Nucleic Acids and Molecular Biology 34

Jorge Cruz-Reyes · Michael W. Gray
Editors

RNA Metabolism in Mitochondria

Springer

Nucleic Acids and Molecular Biology

Volume 34

Series Editor
Allen W. Nicholson
Philadelphia, PA, USA

More information about this series at http://www.springer.com/series/881

Jorge Cruz-Reyes • Michael W. Gray
Editors

RNA Metabolism in Mitochondria

Springer

Editors
Jorge Cruz-Reyes
Department of Biochemistry and Biophysics
Texas A&M University
College Station, Texas, USA

Michael W. Gray
Department of Biochemistry and Molecular Biology, Centre for Comparative Genomics and Evolutionary Bioinformatics
Dalhousie University
Halifax, Nova Scotia, Canada

ISSN 0933-1891 ISSN 1869-2486 (electronic)
Nucleic Acids and Molecular Biology
ISBN 978-3-319-78189-1 ISBN 978-3-319-78190-7 (eBook)
https://doi.org/10.1007/978-3-319-78190-7

Library of Congress Control Number: 2018944140

This Springer imprint is published by the registered company Springer International Publishing AG part of Springer Nature.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

According to endosymbiont theory, mitochondria are the descendants of a once free-living bacterial lineage, phylogenetically related to a specific bacterial group, the α-Proteobacteria. In essence, the mitochondrion may be regarded as a stripped-down and highly re-tailored bacterium. Although initially, in the evolving mitochondrion, gene expression would have been carried out by bacterial-type systems of transcription, posttranscriptional processing, and translation, these systems have been substantially altered during evolutionary transformation of endosymbiont to organelle: and differentially so, in various eukaryotic lineages. In particular, nucleus-encoded proteins, some of which replaced their mitochondrion-encoded counterparts, were recruited to participate in mitochondrial gene expression, imposing a new level of complexity and regulation. In addition, as evolution of the mitochondrial genome took radically different—and in some cases quite bizarre—turns in different eukaryotes, new processes such as *trans*-splicing and RNA editing were added along the pathway of genetic information transfer. Again, this novelty and complexity is strongly lineage specific, making it impossible to develop a unified model of mitochondrial gene expression that is applicable throughout eukaryotes. Comparative analysis of mitochondrial gene expression, a theme of this book, continues to turn up new examples of unusual pathways and processes, particularly among eukaryotic microbes (protists), the least well-studied eukaryotes in this regard.

In this book, we have sought to bring together research using well-developed animal and yeast model systems and focused on a detailed understanding of particular aspects of mitochondrial RNA metabolism, with lineage-specific and novel examples that serve to characterize mitochondria as virtual "playgrounds" in the evolution of new modes of RNA biochemistry and molecular biology. Over the past decade, in particular, great progress has been made in identifying nucleus-encoded protein factors that act in the various pathways of mitochondrial RNA metabolism described herein. Being able to gather together this information in a single volume was a motivating force in developing this book.

In Chap. 1, Drakulic et al. summarize current knowledge about the mitochondrial transcription system in several animal and fungal systems. With the exception of a

single protist lineage (the jakobid flagellates), mitochondrial transcription in all other eukaryotes is under the control of a nucleus-encoded, single-subunit, T3/T7-like bacteriophage RNA polymerase, rather than a mitochondrion-encoded, multi-subunit, bacterial-type enzyme. In Chap. 2, Freyer et al. discuss mitochondrial RNA turnover in Metazoa, highlighting our current understanding of mechanisms to process mitochondrial primary transcripts to yield mature mRNA, rRNA, and tRNA species. This processing involves a variety of nucleus-encoded proteins such as endo- and exonucleases, posttranscriptional modification enzymes, and poly (A) polymerase. Chapter 3, by Saoji and Cox, focuses on the mitochondrial RNase P complex in animals. This universal enzyme forms the 5′ terminus of tRNAs, and here again, the mitochondrial system (an all-protein one) has departed radically from its bacterial counterpart (a ribonucleoprotein complex). In Chap. 4, Weber-Lotfi and Dietrich provide a detailed account of research into mitochondrial RNA trafficking, including the controversial area of RNA import into and export from mitochondria. RNA trafficking is increasingly recognized as an important contributor to function and communication within mitochondria, and between mitochondria and the nucleus, with noncoding RNAs generated in both nucleus and mitochondrion now being implicated.

The final five chapters deal with lineage-specific examples of novel forms of mitochondrial RNA metabolism, four of which are limited to protists. In Chap. 5, Cruz-Reyes et al. discuss U insertion/deletion editing in the kinetoplastid protozoa, a phenomenon discovered in 1986 that provided the first example of what has come to be called “RNA editing”, a now-generic term that encompasses a wide range of biochemically diverse and phylogenetically unrelated mechanisms. In this particular system, editing is carried out by a multi-protein, nucleus-encoded editing complex, the editosome, working in concert with mitochondrion-encoded antisense guide RNAs that provide the nucleotide specificity for editing of primary mRNA transcripts. Chapter 6, by Faktorová et al., details our understanding of a more recently discovered system, in the mitochondria of diplonemids, a group phylogenetically related to kinetoplastids. In the diplonemid case, generation of translatable mRNAs requires extensive *trans*-splicing of transcripts encoded by highly fragmented genes, as well as several different types of RNA editing. In Chap. 7, Dodbele et al. discuss a tRNA editing system that re-tailors the 5′ ends of the acceptor stems of mitochondrial tRNAs, which as encoded in the mitochondrial genome are predicted to contain mismatches in one or more of the first three 5′ positions of the tRNA. This system was initially discovered in several amoebozoan lineages but is now recognized to be more generally distributed among protists and in some fungi. Recent work has identified the proteins that are responsible for this editing. In Chap. 8, Houtz et al. discuss RNA editing in the slime mold, *Physarum polycephalum*. As an amoebozoan, *Physarum* carries out the sort of mitochondrial tRNA editing described in Chap. 7 but has a more extensive insertion/deletion editing system that re-tailors rRNA and tRNA as well as mRNA transcripts—a type of editing so far unique to this protist and its close relatives, in addition to carrying out more limited substitution-type editing of a sort described in other eukaryotes. Finally, in Chap. 9, Takenaka et al. discuss the extensive C-to-U and U-to-C RNA editing that occurs in land plant

mitochondria. Somewhat surprisingly, it turned out that in this system, selection of nucleotides to be edited does not rely on complementary base pairing specified by antisense guide-type RNAs, as in kinetoplastid mitochondria, but on multiple pentatricopeptide repeat (PPR) proteins. This chapter outlines the role played by PPR proteins in this type of editing, as well as the involvement of numerous other non-PPR proteins discovered over the past decade.

The continual elucidation of various factors (both RNA and protein) mediating mitochondrial RNA metabolism has led to increasing recognition of the importance of these factors in overall cellular function. In humans, a number of genetic disorders have now been associated with mutations in mitochondrial RNA metabolism proteins, as several chapters in this book discuss. Continued work in this area will undoubtedly add greatly to our knowledge of overall mitochondrial RNA biology, which in turn will inform our understanding of the functional importance of these processes. As well, continued comparative exploration of the field through extension to new eukaryotic lineages should eventually tell us how much (or how little) of the pathways of mitochondrial RNA metabolism worked out in model systems are generally applicable through the domain Eukarya and perhaps also when and how variations on the general theme emerged during the evolution of the mitochondrion.

Our thanks to the authors who contributed to this book. We learned a lot, and we hope that readers of this volume will, too.

College Station, TX, USA Jorge Cruz-Reyes
Halifax, Nova Scotia, Canada Michael W. Gray

Contents

Chapter 1
The Mitochondrial Transcription Machinery

Srdja Drakulic, Jorge Cuellar, and Rui Sousa

Abstract The involvement of mitochondria in multiple cellular functions beyond generation of ATP creates a need to organize mitochondrial DNA and regulate transcription of mitochondrial genes. The mitochondrial transcription apparatus itself is encoded in the nuclear genome. The central component of this apparatus—the mitochondrial RNA polymerase (mtRNAP)—is homologous to the single-subunit RNAPs encoded by multiple bacteriophages, most notably the well-characterized RNAP encoded by the T7 bacteriophage. Biophysical and biochemical studies have revealed that structure-mechanism relationships are remarkably well-conserved between the phage and mitochondrial RNAPs, with homologous elements in both polymerase classes playing similar roles in promoter recognition, bending, melting, and transcription initiation. However, mtRNAPs are distinct from phage RNAPs, because mtRNAPs in isolation assume a "clenched" conformation in which the large DNA-binding cleft of the polymerase is occluded and other parts of the polymerase involved in promoter binding are sequestered by intramolecular interactions. Interactions between the mtRNAP and mitochondrial transcription factors alter mtRNAP structure to relieve this intramolecular sequestration and unlock the promoter-specific binding and transcriptional activity of the polymerase. There is one such factor required for mitochondrial transcription initiation in yeast and two required factors in mammalian mitochondria, which may allow for greater scope in regulation in higher vs. lower eukaryotes. Thus, mitochondrial transcription relies on an RNAP that is homologous to the phage RNAPs that can function without any accessory factors but exhibits features analogous to nuclear or bacterial transcription

S. Drakulic · J. Cuellar · R. Sousa (✉)
Department for Macromolecular Structures, Centro Nacional de Biotecnología (CNB-CSIC), Madrid, Spain

Department of Biochemistry and Structural Biology, University of Texas Health Science Center, San Antonio, TX, USA
e-mail: SOUSA@uthscsa.edu

J. Cruz-Reyes, M. Gray (eds.), *RNA Metabolism in Mitochondria*, Nucleic Acids and Molecular Biology 34, https://doi.org/10.1007/978-3-319-78190-7_1

in that it requires additional factors to specifically initiate transcription at mitochondrial promoters. In this review, we aim to provide a comprehensive description of the general, common mitochondrial transcription mechanisms and of the variations in these transcription systems, from *Saccharomyces cerevisiae* to *Homo sapiens.*

1.1 Mitochondria: More Than ATP Production

Mitochondria are usually described as the main "power stations" of eukaryotic cells. However, they are also involved in many essential cellular processes such as cell signaling, production of metabolites, cell growth, cell differentiation, programmed cell death (apoptosis), and innate immunity. Thus, not surprisingly, emerging evidence relates mitochondrial dysfunction with a broad range of human disorders and diseases (Wallace and Chalkia 2013). Since the mitochondrial proteome is encoded by both mitochondrial and nuclear genomes, mitochondrial dysfunctions can arise from mutations of genes encoded either in the mitochondrion or nucleus. The higher mutational rate of the mitochondrial genome, probably arising from the combination of the lower efficiency of mitochondrial DNA repair mechanisms and a more "mutagenic" intra-organellar environment (Haag-Liautard et al. 2008), is counterbalanced by mitochondrial heteroplasmy and, in most cases, high mutational threshold values (Tuppen et al. 2010).

1.2 Organization and Packaging of the Mitochondrial Genome

1.2.1 Variation in Mitochondrial Architecture

The first characterized mitochondrial DNAs (mtDNAs) were of vertebrate (chicken, cow, mouse) origin and, surprisingly, revealed circular, covalently closed molecules of ca. 15 kDa (Yoshida et al. 2002). This observation supported the idea, postulated some 80 years earlier by Richard Altmann (based on general, microscopic mitochondrial morphology), that mitochondria evolved from internalized bacteria (Sagan 1967). While this discovery was important for the acceptance of the endosymbiotic theory for the origin of the eukaryotic cell, it also led to the sustained belief that mitochondrial genomes of all eukaryotic organisms are single circular DNA molecules, with rare exceptions to this rule in the form of linear variants in some Ciliata (*Tetrahymena* and *Paramecium)* (Williamson 2002). However, from the constantly growing number of newly characterized mitochondrial genomes, it is evident that there is extensive heterogeneity in the size, content, and architecture of mitochondrial genomes, rather than one general organization (Fig. 1.1a).

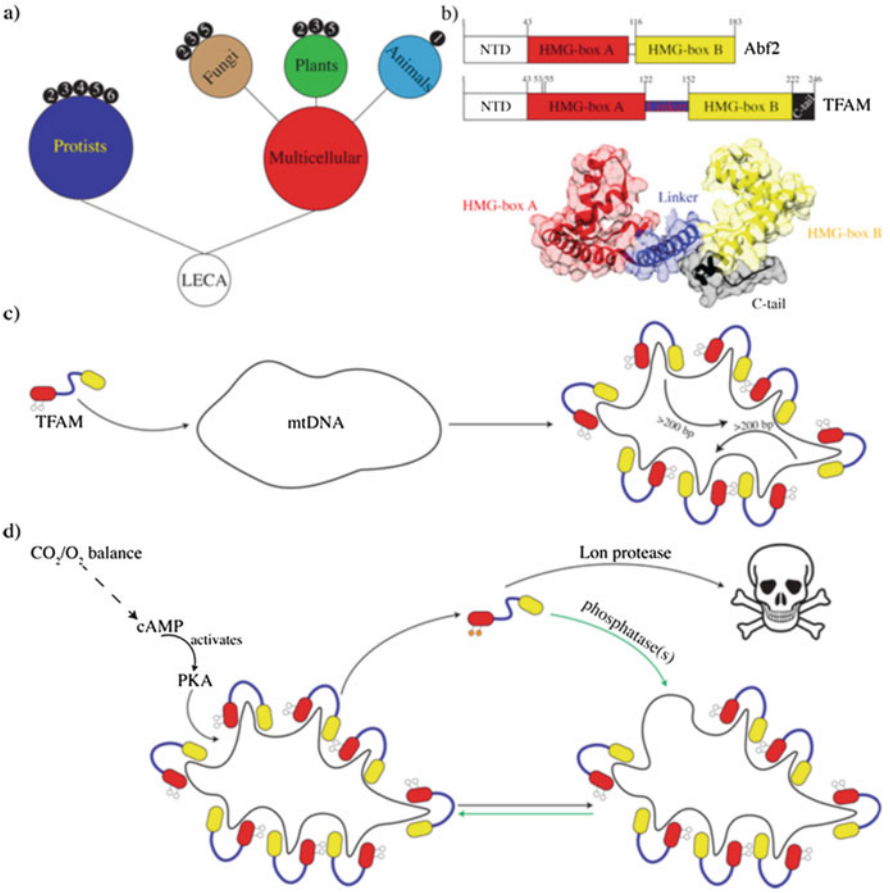

Fig. 1.1 Organization of mitochondrial genomes. (**a**) Distribution of the different types of mitochondrial genomes in eukaryotes. Currently, six main types of mitochondrial genomes can be distinguished according to their size, form, and distribution: (1) single-component circular molecules with sizes from 11 or 12 kbp up to 28 kbp (these are "classical" mitochondrial genomes), (2) single-component circular molecules with sizes from 22 kbp up to 1 Mbp, (3) circular molecules of more than 22 kbp with the simultaneous presence of plasmid-like molecules, (4) multipartite genomes comprised of heterogeneous populations of circular molecules, (5) single-component linear molecules, and (6) multipartite linear molecules comprised of heterogeneous populations of linear molecules. LECA is the last eukaryotic common ancestor from which diverged both uni- and multicellular eukaryotic organisms. (**b**) Schematic comparison of yeast Abf2 and human TFAM sequences. The color code of the structural domains is maintained for the atomic structure of TFAM, pdb#: 3TMM. (**c**) The human mitochondrial genome is present in the form of a single-component circular molecule of ~16.8 kDa packaged into ~100-nm nucleotides. The first level of compaction is obtained through the bending of mtDNA by TFAM molecules, while further compaction is achieved by dimerization of TFAMs and subsequent looping. (**d**) TFAM's affinity for DNA can be modulated in response to the metabolic/energy requirements of the cell through phosphorylation of serine residues at positions 53 and 56 (depicted as circles). The unbound TFAM may either be degraded by Lon protease or recycled by the activity of mitochondrial phosphatases

1.2.2 Compaction of Mitochondrial DNA

Mitochondrial DNAs (mtDNAs) are not present in "naked" or unprotected form but are instead assembled into dynamic nucleoprotein complexes—nucleoids—which allow effective genome maintenance and integration into cellular signaling (Gilkerson et al. 2013). The principal components of mitochondrial yeast and mammalian nucleoids are, respectively, Abf2 (ARS2-binding factor) and TFAM (transcription factor A, mitochondrial) (Diffley and Stillman 1991; Parisi and Clayton 1991). These two homologous proteins belong to the high-mobility group (HMG) of nonhistone DNA-binding proteins (Chen and Butow 2005). Independent characterizations of Abf2/DNA (Friddle et al. 2004) and TFAM/DNA complexes (Ngo et al. 2014; Rubio-Cosials et al. 2011) led to the conclusion that these two proteins reduce the length of extended mtDNA via a similar mechanism, by inducing local U-turns in the DNA. In vitro, both proteins bind DNA in a non-sequence-specific manner, with a similar occupancy rate of one protein for every 20–30 bp of DNA (Diffley and Stillman 1992; Farge et al. 2012). However, data suggest that in mammalian tissues and cells (Alam et al. 2003; Ekstrand et al. 2004; Kukat et al. 2011), TFAM is present in sufficient amounts to fully coat the mitochondrial genome (Ngo et al. 2014), while physiological amounts of Abf2 in yeast limit its compacting capacity (Chen and Butow 2005). The latter suggests the existence of an additional packaging mechanism in yeast mitochondria. Crystal structures of TFAM/DNA complexes (Ngo et al. 2011, 2014; Rubio-Cosials et al. 2011) reveal the two HMG-box domains, HMG-A and HMG-B, wedged into the minor groove of mtDNA, while a helical linker stabilizes the entire structure through electrostatic interactions with the negatively charged backbone of the DNA (Fig. 1.1b). It was also shown that TFAM monomers tend to form dimers in the presence of DNA, which drives looping of the DNA as an additional level of packaging. The dimerization is constrained to TFAM monomers that are separated by more than 200 bp (Fig. 1.1c), but additional constraints on dimerization may limit the number of loops that form in the mitochondrial DNA (Ngo et al. 2014).

1.3 The Mitochondrial Transcription Machinery

1.3.1 Mitochondrial RNAP Is Homologous to Single-Subunit Phage RNAPs but Uses Accessory Transcription Factors Analogous to Nuclear Transcription

During the evolution of the eukaryotic cell, the genomic content of the bacterial endosymbiont that was the mitochondrial ancestor has been drastically reduced, both by gene loss and by transfer of genes from the mitochondrion to the nucleus. Only a small fraction of the original endosymbiont genes remain in the mitochondrial genome. Some of the genes expressed in the mitochondrial proteome derive from

neither the original endosymbiont nor its host but have been acquired from other sources. Examples of this include the mitochondrial RNA polymerase (mtRNAP), which is of bacteriophage origin (exceptions can be found among jakobid protists, which have been reported to contain a multi-subunit mitochondrial RNA polymerase (Burger et al. 2013)). However, the mitochondrial transcription machinery is not as simple as that of the bacteriophage that uses a homologous single-subunit RNAP to transcribe their genomes. Instead, mitochondrial transcription also exhibits features analogous to transcription by multi-subunit RNAP families, as mtRNAPs, unlike most single-subunit phage RNAPs, require auxiliary factor(s) for initiation of transcription. This requirement potentially allows for additional levels of regulation of mitochondrial transcription.

1.3.2 The Mammalian (Human) Mitochondrial Transcription Machinery

1.3.2.1 Mammalian Mitochondria Utilize Three Proteins for Initiation of Transcription

The mammalian mitochondrial transcription apparatus is a three-component system composed of POLRMT (mitochondrial RNA polymerase), TFB2M (transcription factor B2, mitochondrial), and TFAM (the homologue of yeast Abf2). Estimates of the number of TFAM molecules per molecule of mtDNA and of TFAM's affinity for mitochondrial promoter DNA imply that the light-strand promoter (LSP) and the heavy-strand promoter 1 (HSP1) are fully occupied by TFAM prior to assembly of the complete transcription initiation complex (IC) (Kukat et al. 2011; Ramachandran et al. 2016). FRET studies have revealed multiple TFAMs bound in a sequence-specific manner to bent, upstream LSP elements and, in a non-sequence-specific manner, to DNA downstream of the LSP (Ramachandran et al. 2016). As the K_d values of TFAM/LSP and POLRMT/TFB2M/LSP interactions are similar (7 and 5 nM, respectively) (Ramachandran et al. 2016), it is possible that the latter complex displaces the downstream-bound TFAM and interacts with the TFAM located at the upstream LSP element. Photo-cross-linking experiments have revealed multiple interaction sites between the C-terminal tail of TFAM and the N-terminal region of POLRMT (Table 1.1) (Morozov et al. 2014, 2015).

1.3.2.2 Sequence-Specific Binding of TFAM to Human Mitochondrial Promoters

In addition to the non-sequence-specific binding mode that is essential for TFAM compaction and organization of mitochondrial DNA into nucleoid structures, TFAM also binds DNA in a sequence-specific manner that facilitates promoter recognition and assembly of the mitochondrial transcriptional machinery (Chen and Butow

Table 1.1 Mapped interactions between TFAM or TFB2M and POLRMT

TFAM 217	POLRMT 120–134
TFAM 227	POLRMT 120–150
TFAM 228	POLRMT 444–462 (D-helix)
TFAM 233	POLRMT 444–473 (D-helix)
TFB2M 315–352 (α8-helix)	POLRMT 588–604 (B-loop)

On the left are the TFAM or TFB2M residues that have been substituted with the pBpa (parabenzoyl phenylalanine) cross-linker (Morozov et al. 2014, 2015), while on the right are the regions of POLRMT that have been mapped by hydroxylamine digestion and cyanogen bromide (CNBr) as interacting with the transcription factor residues (Morozov et al. 2015)

2005). Specifically, human mtDNA contains three promoters, the light-strand promoter (LSP) and two heavy-strand promoters (HSP1 and HSP2), located in the displacement-loop (D-loop) region of the genome (Chang and Clayton 1984; Fisher and Clayton 1985). Biochemical and structural studies indicate that the HMG-A domain binds promoter DNA first (Gangelhoff et al. 2009; Ngo et al. 2011, 2014; Rubio-Cosials et al. 2011). This induces conformational changes both in TFAM and DNA, allowing the lower-affinity HMG-B domain to bind (Lu et al. 2013). LSP and HSP1 contain both distal and proximal binding sites for the two HMG domains of TFAM, and TFAM binds these elements in a specific orientation (Fisher et al. 1987; Ngo et al. 2014). This is the basis for the differential recognition and transcriptional activity of each promoter. TFAM exhibits lower affinity for HSP1 than for the LSP (Kanki et al. 2004; Ngo et al. 2014). On HSP1, HMG-B domain binds to the proximal binding element, resulting in loss of additional contacts observed in the TFAM/LSP complex. In addition, the positioning of the TFAM C-terminal tail, which is essential for interaction with POLRMT (human mitochondrial RNA polymerase) and transcription initiation (Dairaghi et al. 1995), abolishes the requirement for mtDNA bending for initiation of transcription (Ngo et al. 2014). The absence of a corresponding C-terminal tail in Abf2 may explain why it plays no role in transcription initiation in *S. cerevisiae* (Fig. 1.1b) (Dairaghi et al. 1995). On the LSP, the HMG-B domain binds to a distal element, which has two consequences. First, the different binding orientation allows TFAM to make more contacts with the LSP than with HSP1, leading to the increased stability of the LSP/TFAM complex. Second, the TFAM C-terminal tail is placed further from the initiation start site on the light strand. This creates a requirement for TFAM to impose bending on the DNA so as to bring the transcription machinery in close proximity to the LSP transcription start site (Ngo et al. 2011).

1.3.2.3 Phosphorylation May Regulate TFAM Binding to Promoter DNA

Since the initial contacts between mtDNA and TFAM involve the HMG-A domain, the latter's binding affinity would most likely be an important regulatory target

(Lu et al. 2013). Protein kinase A (PKA)-mediated phosphorylation of serines at positions 55 and 56 causes TFAM to be released from DNA and makes it prone to Lon-mediated proteolysis (Fig. 1.1d) (Lu et al. 2013), though it is also possible that DNA-free TFAM is not always degraded but can be dephosphorylated and stored within nucleoids until its use for genome coating (Fig. 1.1d).

There is evidence that cAMP-mediated activation of PKA and subsequent TFAM phosphorylation may occur in response to changes in metabolically generated CO_2/HCO_3^- (as may occur during ischemia) (Acin-Perez et al. 2009). And in *Drosophila*, it has been shown that the prune (Pn) phosphodiesterase downregulates cAMP levels in the mitochondrial matrix and promotes mtDNA replication by stabilizing TFAM. There may therefore be multiple links between cAMP signaling, TFAM, and mitochondrial transcription (Fig. 1.1d).

1.3.2.4 Additional TFAM Functions

Because transcripts obtained from transcription of the light chain are required for mtDNA replication and because mtDNA stability is enhanced through packaging into nucleoids, TFAM is important for the maintenance and regulation of mtDNA copy number (Campbell et al. 2012; Wen et al. 2016; Zhang et al. 2015). A potential mtDNA repair function has been also proposed for TFAM, as its affinity for oxidatively or cisplatin-damaged DNA is greater than for undamaged DNA (Yoshida et al. 2002, 2003).

1.3.2.5 TFB2M and POLRMT Bind the Promoter After TFAM

Though TFAM binding is the first step in formation of the transcription complex, it is unclear whether POLRMT subsequently binds alone, followed by TFB2M, or whether POLRMT and TFB2M bind together as a complex. The latter would be consistent with observations that POLRMT/TFB2M exhibits almost 1.5- to 2-fold higher stability on LSP than POLRMT alone (Ramachandran et al. 2016) and that POLRMT/TFB2M forms a complex on DNA that is stable enough to survive gel filtration (Yakubovskaya et al. 2014). Thus, the question whether POLRMT alone or TFB2M-bound POLRMT interacts with TFAM and LSP DNA remains to be settled. Once all three proteins are bound to the LSP, the promoter is melted, and the open complex (OC) forms.

Insights into the promoter recognition and melting mechanisms of the mitochondrial transcription machinery emerge from studies of the extensively studied and homologous T7 phage RNAP. The latter employs three structural motifs for promoter binding (formation of closed complex) and melting (transition to the open complex, OC): the AT-rich recognition loop (residues 93–101), the specificity loop (residues 739–770), and the intercalating β-hairpin (residues 230–245) (Fig. 1.2a). The interaction of the AT-rich recognition loop with the region that extends from 13 to 17 base pairs upstream of the transcription start site induces structural

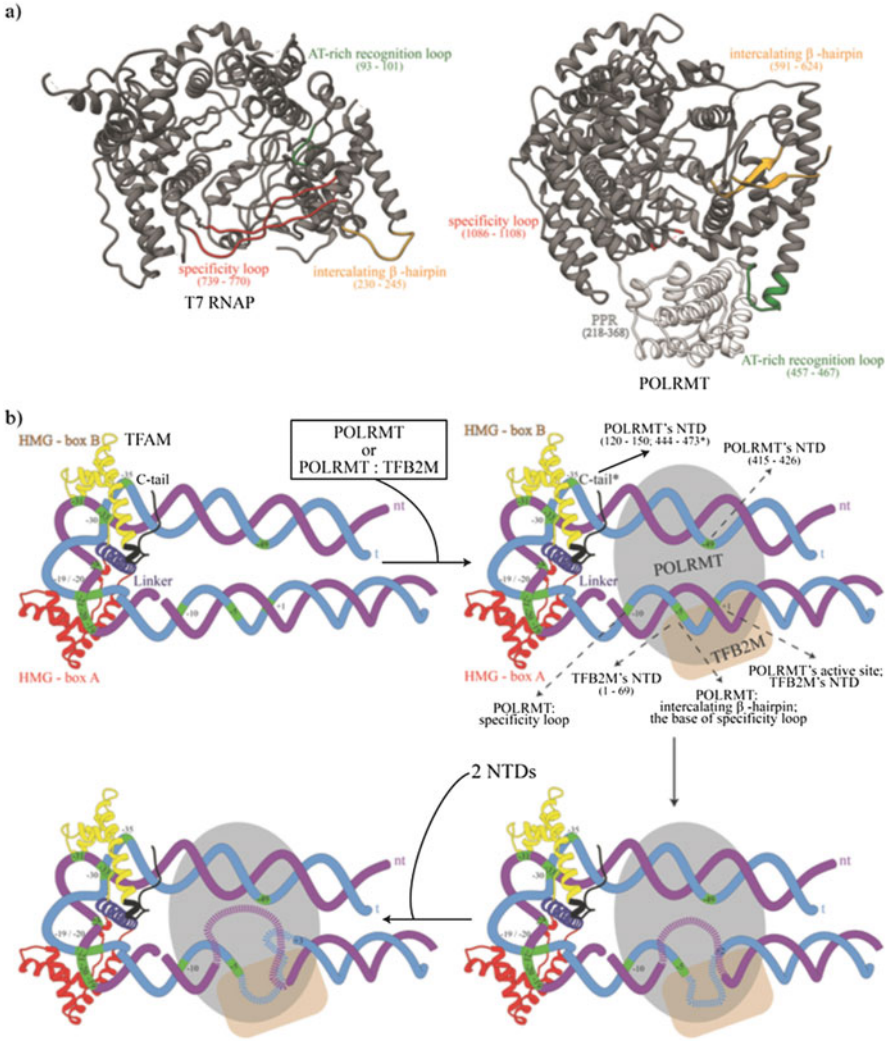

Fig. 1.2 Mechanism of the initial steps of transcription initiation by the human mitochondrial machinery. (**a**) The atomic structures of T7RNAP (pdb 4RNP) and POLRMT (pdb 3SPA). The structural elements of T7RNAP responsible for promoter binding, bending, and melting and their analogues in POLRMT—the specificity loop, intercalating β-hairpin, and AT-rich recognition loop—have been colored in red, orange, and green, respectively. (**b**) TFAM recruits POLRMT or POLRMT/TFB2M complexes to the LSP through interactions between TFAM's C-terminal tail and POLRMT's NTD. Specific interaction sites between TFAM and POLRMT are presented in Table 1.1. The presence of both transcription factors induces conformational changes in POLRMT that drive the transition from the clenched to the open conformation and release of NTE's sequestration of the AT-rich recognition loop. These structural rearrangements also involve the intercalating β-hairpin that is pushed toward the DNA duplex at the −5 position. The newly formed interactions with promoter DNA are additionally reinforced by TFB2M's NTD (residues 1–69), resulting in the increased stability of the IC. The bending of the promoter leads to disruption of the

reorganization in both T7 RNAP and the promoter DNA. This allows the correct positioning of the remaining motifs and corresponding promoter elements: the specificity loop with the region 7 to 11 bp upstream of the start site and the intercalating hairpin, which inserts between the template and non-template strand immediately upstream of the start site, results in an initial open complex with a melted region extending from −4 to +1 (Cheetham et al. 1999).

The POLRMT shows conservation of all three of these phage RNAP structural elements that are involved in promoter recognition and melting. POLRMT contains a large NH_2-terminal domain (NTD, residues 1–647) that exhibits structural similarity to that of T7 RNAP, despite the absence of identifiable similarity at the sequence level. Structural elements corresponding to the T7 RNAP AT-rich recognition loop (residues 457–467) and intercalating β-hairpin (residues 591–624) are located in the NTD, while the corresponding specificity loop (residues 1086–1108) is within the C-terminal domain (CTD) (Fig. 1.2a). In the absence of TFB2M and TFAM, POLRMT's NH_2-terminal extension (NTE, residues 42–368) sequesters the AT-rich loop, while the CTD is in a "clenched" conformation that partially closes its deep DNA-binding cleft (Ringel et al. 2011). These are the structural mechanisms that account for POLRMT's observed inability to initiate transcription from native promoters (Ramachandran et al. 2016).

The interaction of TFAM's C-terminal tail and TFB2M's α8-helix with POLRMT's NTD (Table 1.1) is proposed to relieve the sequestration of the AT-rich recognition loop by the NTE and to push the intercalating β-hairpin toward the DNA duplex at the −5 position (Morozov et al. 2015). In the case of T7 RNAP, binding of the AT-rich recognition and specificity loops bends the promoter DNA by 40–60° at the −1 base pair. This bending drives opening of the duplex DNA in the TATA region (−1 to −4). Val235 of the intercalating β-hairpin is placed between the −5 and −4 bps and stabilizes the upstream edge of the newly formed transcription bubble (Cheetham et al. 1999; Cheetham and Steitz 1999). Displacement of POLRMT's intercalating β-hairpin may lead to formation of a pocket between the β-hairpin and the extended, helical "thumb" subdomain of the CTD. This may allow the TFB2M NTE (residues 1–69) to be placed in close proximity of the POLRMT active site (Morozov et al. 2015). Contacts between the TFB2M NTD and the −5 and +1 bases of the template strand (Morozov et al. 2015), the upstream and downstream edges of the transcription bubble, and the initiating ATP (Sologub et al. 2009) all contribute to the role of TFB2M in stabilization of the OC. In addition to being required for formation and stabilization of the OC, TFAM is also required to allow extension of transcripts beyond 2 nts in length (Ramachandran et al. 2016) (Fig. 1.2b).

Fig. 1.2 (continued) duplex from −4 to +1. TFB2M interacts with both the −5 and +1 bases, participating in promoter melting and stabilization of the newly formed transcription bubble. TF2B2M and TFAM play additional roles: interaction with the initial (priming or +1) NTP and successful extension of the nascent RNA beyond 2 nts, respectively

1.3.3 The Yeast (S. cerevisiae) Mitochondrial Transcription Machinery

1.3.3.1 Yeast mtRNAP Uses a Single Transcription Factor to Initiate from the 28 Promoters Found in the Yeast Mitochondrial Genome

The mtRNAP (Rpo41) of the baker's yeast *S. cerevisiae*, unlike its human homologue, requires only one auxiliary factor, Mtf1 (mitochondrial transcription factor 1), for the initial steps of RNA synthesis. This machinery can start transcription from 14 different promoters distributed along the population of polydisperse linear DNA molecules that comprise the yeast mitochondrial genome (Fig. 1.1a) (Turk et al. 2013). Each of these promoters contains a conserved nine-nucleotide consensus sequence ($^{-8}$ATATAAGTA^{+1}). Promoter activity is strongly dependent on the non-template strand sequence at positions +1 and +2, with an A occurring at +1 in all promoters, while the +2 position is more variable (Deshpande and Patel 2012). The strongest promoters contain an A at this position, though promoters with a G are nearly as active, but promoters with a pyrimidine are significantly weaker (Deshpande and Patel 2014).

1.3.3.2 Rpo41 Alone Binds Nonspecifically to DNA

Rpo41 alone binds promoter and non-promoter DNA sequences equally well with a K_d of ~48–66 nM (Deshpande and Patel 2012). This non-sequence-specific binding could allow Rpo41 to bind weakly to mtDNA and slide along it until it encounters a promoter. Cryo-EM structures of Rpo41 bound to DNA alone and together with MTF1 (Drakulic et al. 2014), and crystal structures of human mtRNAP (Ringel et al. 2011), suggest that Rpo41 interacts with mtDNA in a clenched conformation, prior to binding with Mtf1 (Figs. 1.1, 1.2 and 1.3).

1.3.3.3 MTF1 Drives Conformational Changes in Rpo41 that Unlock the Latter's Sequence-Specific DNA-Binding and Promoter-Opening Activities

The inability of Rpo41, like POLRMT, to initiate transcription alone may be due to the polymerase, in the absence of MTFs, assuming the clenched conformation and to the NTE sequestering the AT-rich recognition loop. The presence of Mtf1 dramatically increases the stability of all Rpo41/DNA complexes, but this effect is more pronounced on promoter DNA complexes, as the K_d values for MTF1/Rpo41/promoter DNA complexes are fourfold lower than for non-promoter complexes

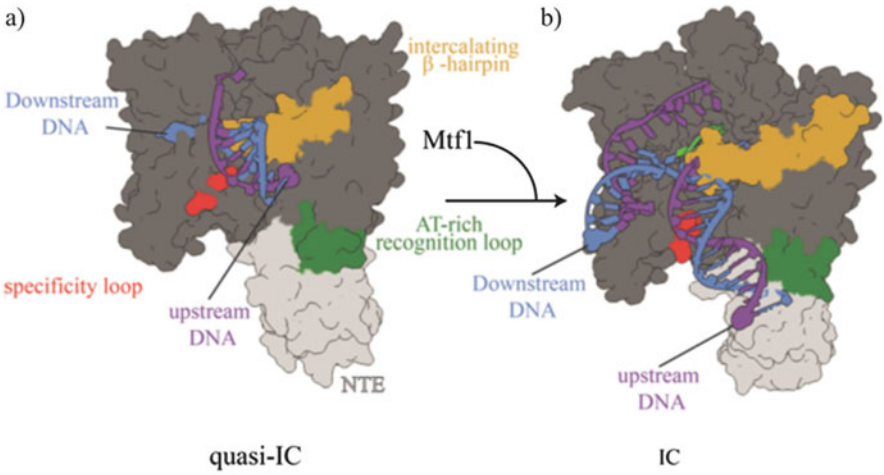

Fig. 1.3 Mechanism of the initial steps of transcription initiation by the yeast mitochondrial machinery. (**a**) Model of Rpo41 alone bound to a yeast mitochondrial promoter. This model is based on cryo-EM images of Rpo41 complexes formed in the absence of MTF1 using a "pre-melted" promoter that contains a heteroduplex in the region that becomes melted in the open complex. Such pre-melted promoters stabilize the association with Rpo41, allowing stable complexes to be formed in the absence of MTF1. The AT-rich recognition loop, specificity loop, intercalating β-hairpin, and PPR region of the NTE are colored green, red, orange, and light gray, respectively. The template and non-template DNA strands and RNA are colored blue, purple, and light green, respectively. In the absence of MTF1, Rpo41 binds in a non-sequence-specific manner, with the AT-rich recognition loop hidden by the NTE and the upstream region of the DNA placed far from the AT-rich recognition loop. (**b**) Cryo-EM-based model of Rpo41/Mtf1/promoter initiation complex (IC). The magenta-colored transparent ellipse shows where the ellipsoid MTF1 molecule is bound. Binding of Mtf1 to the Rpo41/promoter complex brings the upstream end of the DNA in close proximity to the AT-rich recognition loop. This is coupled with opening (unclenching) of the Rpo41 template-binding cleft. In addition, Mtf1 establishes contacts with the promoter DNA and contributes to stabilization of the IC and promoter melting

(Tang et al. 2011). Binding of Mtf1 induces structural rearrangements both in Rpo41 and the promoter. The cryo-EM structures reveal that, at this step, Rpo41 assumes the open conformation similar to POLRMT during elongation (Schwinghammer et al. 2013), with a wider binding cleft that is formed through rotations of the thumb and finger subdomains. During this step, the promoter DNA becomes additionally bent, from 52° to 89° (Tang et al. 2011), which brings the upstream promoter elements closer to the AT-rich recognition loop (Drakulic et al. 2014) (Figs. 1.1, 1.2 and 1.3). The increased stability of the Rpo41/promoter DNA/Mtf1 open complex is derived primarily from these additional contacts between Rpo41 and promoter DNA, together with interactions between Mtf1 and the unwound NT strand (Drakulic et al. 2014; Paratkar and Patel 2010). In addition, the C-terminal tail of Mtf1 is inserted between the melted DNA strands where it may contribute to promoter opening (Drakulic et al. 2014; Savkina et al. 2010).

1.3.4 Mitochondrial Transcription: Variations Within a Common Theme

In little more than a decade, a great deal of progress has been made in understanding mitochondrial transcription mechanisms. However, while the main outlines of the picture are known, important details have not been fully clarified. In part, this reflects the paucity of high-resolution structural data of all of the different steps in the transcription process: from initial transcription factor binding, subsequent recruitment of the polymerase to form a complete IC, promoter bending and melting, initiation of transcription, and extension of initial transcripts to a length that drives promoter release and the transition to elongation, elongation itself, and, ultimately, termination. In the absence of such data, it is tempting to assume that similar mitochondrial transcription mechanisms and machinery occur in all organisms.

This review highlights that this is not the case: while there are many similarities, the mitochondrial transcription machinery differs in yeast and humans. These differences may, in part, be due to the differences in genome structure and organization in fungi and mammals. The necessity for TFAM in human mitochondrial gene transcription, for example, might be a consequence of the circular genome organization and much higher GC content of human mitochondrial promoter regions compared to the promoters found in linear yeast mitochondrial genomes. In addition, the involvement of two, rather than just one, accessory factors in human vs. yeast mitochondrial transcription allows for greater scope and complexity in regulation, which may be important in higher vs. lower eukaryotes.

With our picture of the mechanisms of mitochondrial transcription that is becoming clearer, future studies will need to focus on extending our understanding of how mitochondrial transcription is coupled to and coordinated with other processes. This includes an understanding of the role of mitochondrial transcription in translation and in mitochondrial DNA replication, as well as an understanding of how the expression of nuclear and mitochondrial genes encoding mitochondrial proteins is coordinated.

Acknowledgments We thank C. Mark for the editorial assistance. This work was supported by the Spanish Ministry of Economy (grant BFU2016-75984) and GM118933 (to R.S.).

References

Acin-Perez R, Salazar E, Kamenetsky M, Buck J, Levin LR, Manfredi G (2009) Cyclic AMP produced inside mitochondria regulates oxidative phosphorylation. Cell Metab 9:265–276. https://doi.org/10.1016/j.cmet.2009.01.012

Alam TI et al (2003) Human mitochondrial DNA is packaged with TFAM. Nucleic Acids Res 31:1640–1645

Burger G, Gray MW, Forget L, Lang BF (2013) Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists. Genome Biol Evol 5:418–438. https://doi.org/10.1093/gbe/evt008

Campbell CT, Kolesar JE, Kaufman BA (2012) Mitochondrial transcription factor a regulates mitochondrial transcription initiation, DNA packaging, and genome copy number. Biochim Biophys Acta 1819:921–929. https://doi.org/10.1016/j.bbagrm.2012.03.002

Chang DD, Clayton DA (1984) Precise identification of individual promoters for transcription of each strand of human mitochondrial DNA. Cell 36:635–643

Cheetham GM, Steitz TA (1999) Structure of a transcribing T7 RNA polymerase initiation complex. Science 286:2305–2309

Cheetham GM, Jeruzalmi D, Steitz TA (1999) Structural basis for initiation of transcription from an RNA polymerase-promoter complex. Nature 399:80–83

Chen XJ, Butow RA (2005) The organization and inheritance of the mitochondrial genome. Nat Rev Genet 6:815–825. https://doi.org/10.1038/nrg1708

Dairaghi DJ, Shadel GS, Clayton DA (1995) Addition of a 29 residue carboxyl-terminal tail converts a simple HMG box-containing protein into a transcriptional activator. J Mol Biol 249:11–28. https://doi.org/10.1006/jmbi.1995.9889

Deshpande AP, Patel SS (2012) Mechanism of transcription initiation by the yeast mitochondrial RNA polymerase. Biochim Biophys Acta 1819:930–938. https://doi.org/10.1016/j.bbagrm.2012.02.003

Deshpande AP, Patel SS (2014) Interactions of the yeast mitochondrial RNA polymerase with the +1 and +2 promoter bases dictate transcription initiation efficiency. Nucleic Acids Res 42:11721–11732. https://doi.org/10.1093/nar/gku868

Diffley JF, Stillman B (1991) A close relative of the nuclear, chromosomal high-mobility group protein HMG1 in yeast mitochondria. Proc Natl Acad Sci U S A 88:7864–7868

Diffley JF, Stillman B (1992) DNA binding properties of an HMG1-related protein from yeast mitochondria. J Biol Chem 267:3368–3374

Drakulic S et al (2014) Yeast mitochondrial RNAP conformational changes are regulated by interactions with the mitochondrial transcription factor. Nucleic Acids Res 42:11246–11260. https://doi.org/10.1093/nar/gku795

Ekstrand MI et al (2004) Mitochondrial transcription factor a regulates mtDNA copy number in mammals. Hum Mol Genet 13:935–944. https://doi.org/10.1093/hmg/ddh109

Farge G et al (2012) Protein sliding and DNA denaturation are essential for DNA organization by human mitochondrial transcription factor a. Nat Commun 3:1013. https://doi.org/10.1038/ncomms2001

Fisher RP, Clayton DA (1985) A transcription factor required for promoter recognition by human mitochondrial RNA polymerase. Accurate initiation at the heavy- and light-strand promoters dissected and reconstituted in vitro. J Biol Chem 260:11330–11338

Fisher RP, Topper JN, Clayton DA (1987) Promoter selection in human mitochondria involves binding of a transcription factor to orientation-independent upstream regulatory elements. Cell 50:247–258

Friddle RW et al (2004) Mechanism of DNA compaction by yeast mitochondrial protein Abf2p. Biophys J 86:1632–1639. https://doi.org/10.1016/S0006-3495(04)74231-9

Gangelhoff TA, Mungalachetty PS, Nix JC, Churchill ME (2009) Structural analysis and DNA binding of the HMG domains of the human mitochondrial transcription factor a. Nucleic Acids Res 37:3153–3164. https://doi.org/10.1093/nar/gkp157

Gilkerson R, Bravo L, Garcia I, Gaytan N, Herrera A, Maldonado A, Quintanilla B (2013) The mitochondrial nucleoid: integrating mitochondrial DNA into cellular homeostasis. Cold Spring Harb Perspect Biol 5:a011080. https://doi.org/10.1101/cshperspect.a011080

Haag-Liautard C, Coffey N, Houle D, Lynch M, Charlesworth B, Keightley PD (2008) Direct estimation of the mitochondrial DNA mutation rate in *Drosophila melanogaster*. PLoS Biol 6:e204. https://doi.org/10.1371/journal.pbio.0060204

Kanki T et al (2004) Architectural role of mitochondrial transcription factor a in maintenance of human mitochondrial DNA. Mol Cell Biol 24:9823–9834. https://doi.org/10.1128/MCB.24.22.9823-9834.2004

Kukat C, Wurm CA, Spahr H, Falkenberg M, Larsson NG, Jakobs S (2011) Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA. Proc Natl Acad Sci U S A 108:13534–13539. https://doi.org/10.1073/pnas.1109263108

Lu B et al (2013) Phosphorylation of human TFAM in mitochondria impairs DNA binding and promotes degradation by the AAA+ Lon protease. Mol Cell 49:121–132. https://doi.org/10.1016/j.molcel.2012.10.023

Morozov YI, Agaronyan K, Cheung AC, Anikin M, Cramer P, Temiakov D (2014) A novel intermediate in transcription initiation by human mitochondrial RNA polymerase. Nucleic Acids Res 42:3884–3893. https://doi.org/10.1093/nar/gkt1356

Morozov YI, Parshin AV, Agaronyan K, Cheung AC, Anikin M, Cramer P, Temiakov D (2015) A model for transcription initiation in human mitochondria. Nucleic Acids Res 43:3726–3735. https://doi.org/10.1093/nar/gkv235

Ngo HB, Kaiser JT, Chan DC (2011) The mitochondrial transcription and packaging factor Tfam imposes a U-turn on mitochondrial DNA. Nat Struct Mol Biol 18:1290–1296. https://doi.org/10.1038/nsmb.2159

Ngo HB, Lovely GA, Phillips R, Chan DC (2014) Distinct structural features of TFAM drive mitochondrial DNA packaging versus transcriptional activation. Nat Commun 5:3077. https://doi.org/10.1038/ncomms4077

Paratkar S, Patel SS (2010) Mitochondrial transcription factor Mtf1 traps the unwound non-template strand to facilitate open complex formation. J Biol Chem 285:3949–3956. https://doi.org/10.1074/jbc.M109.050732

Parisi MA, Clayton DA (1991) Similarity of human mitochondrial transcription factor 1 to high mobility group proteins. Science 252:965–969

Ramachandran A, Basu U, Sultana S, Nandakumar D, Patel SS (2016) Human mitochondrial transcription factors TFAM and TFB2M work synergistically in promoter melting during transcription initiation. Nucleic Acids Res. https://doi.org/10.1093/nar/gkw1157

Ringel R, Sologub M, Morozov YI, Litonin D, Cramer P, Temiakov D (2011) Structure of human mitochondrial RNA polymerase. Nature 478:269–273. https://doi.org/10.1038/nature10435

Rubio-Cosials A et al (2011) Human mitochondrial transcription factor a induces a U-turn structure in the light strand promoter. Nat Struct Mol Biol 18:1281–1289. https://doi.org/10.1038/nsmb.2160

Sagan L (1967) On the origin of mitosing cells. J Theor Biol 14:255–274

Savkina M, Temiakov D, McAllister WT, Anikin M (2010) Multiple functions of yeast mitochondrial transcription factor Mtf1p during initiation. J Biol Chem 285:3957–3964. https://doi.org/10.1074/jbc.M109.051003

Schwinghammer K, Cheung AC, Morozov YI, Agaronyan K, Temiakov D, Cramer P (2013) Structure of human mitochondrial RNA polymerase elongation complex. Nat Struct Mol Biol 20:1298–1303. https://doi.org/10.1038/nsmb.2683

Sologub M, Litonin D, Anikin M, Mustaev A, Temiakov D (2009) TFB2 is a transient component of the catalytic site of the human mitochondrial RNA polymerase. Cell 139:934–944

Tang GQ, Deshpande AP, Patel SS (2011) Transcription factor-dependent DNA bending governs promoter recognition by the mitochondrial RNA polymerase. J Biol Chem 286:38805–38813. https://doi.org/10.1074/jbc.M111.261966

Tuppen HA, Blakely EL, Turnbull DM, Taylor RW (2010) Mitochondrial DNA mutations and human disease. Biochim Biophys Acta 1797:113–128. https://doi.org/10.1016/j.bbabio.2009.09.005

Turk EM, Das V, Seibert RD, Andrulis ED (2013) The mitochondrial RNA landscape of *Saccharomyces cerevisiae*. PLoS One 8:e78105. https://doi.org/10.1371/journal.pone.0078105

Wallace DC, Chalkia D (2013) Mitochondrial DNA genetics and the heteroplasmy conundrum in evolution and disease. Cold Spring Harb Perspect Biol 5:a021220. https://doi.org/10.1101/cshperspect.a021220

Wen S, Gao J, Zhang L, Zhou H, Fang D, Feng S (2016) p53 increase mitochondrial copy number via up-regulation of mitochondrial transcription factor a in colorectal cancer. Oncotarget. https://doi.org/10.18632/oncotarget.12514

Williamson D (2002) The curious history of yeast mitochondrial DNA. Nat Rev Genet 3:475–481. https://doi.org/10.1038/nrg814

Yakubovskaya E et al (2014) Organization of the human mitochondrial transcription initiation complex. Nucleic Acids Res 42:4100–4112. https://doi.org/10.1093/nar/gkt1360

Yoshida Y et al (2002) Human mitochondrial transcription factor a binds preferentially to oxidatively damaged DNA. Biochem Biophys Res Commun 295:945–951

Yoshida Y, Izumi H, Torigoe T, Ishiguchi H, Itoh H, Kang D, Kohno K (2003) P53 physically interacts with mitochondrial transcription factor A and differentially regulates binding to damaged DNA. Cancer Res 63:3729–3734

Zhang F, Qi Y, Zhou K, Zhang G, Linask K, Xu H (2015) The cAMP phosphodiesterase prune localizes to the mitochondrial matrix and promotes mtDNA replication by stabilizing TFAM. EMBO Rep 16:520–527. https://doi.org/10.15252/embr.201439636

Chapter 2
Mitochondrial RNA Turnover in Metazoa

Christoph Freyer, Paula Clemente, and Anna Wredenberg

Abstract Correct regulation of mitochondrial gene expression is central to controlling mitochondrial function, and defects in all aspects of gene expression have been observed in a range of disorders. Many of the central mechanisms involved are not yet understood, but high conservation among many species allows for the use of a range of model systems to further our understanding of mitochondrial gene expression. Studies from mice and fruit flies have shown functional conservation with the human system and have provided important insights into general mechanisms inside the mitochondrial network. Here we describe recent insights into mitochondrial gene expression, focusing on observations made in the fruit fly, *Drosophila melanogaster*.

The engulfing of an α-proteobacterium by another cell (perhaps an archaebacterium) some 2 billion years ago gave rise to the first eukaryotic cell, and this endosymbiont, which evolved to become the mitochondrion, has since risen to be a central hub in cellular metabolism. This endosymbiosis has been suggested to dramatically increase the available energy to the newly formed composite cell through sharing of common requirements and reduction of the endosymbiont's genome (Lane and Martin 2010; Lane 2017). This increase in available energy ultimately allowed for the increased complexity seen in today's eukaryotic cells, with highly specialised compartmentalisation and restructuring of the two genomes. In this scenario, mitochondria, through their bioenergetic membranes, provided the host cell with energy and essential metabolites. However, it seems that bioenergetic membranes require tight control of the cellular redox state (Allen 2017) and large bacterial cells have therefore multiple genomes attached to their plasma membranes. The bacterial ancestry of mitochondria reflects this characteristic, with the majority of mitochondria retaining a high degree of polyploidy, with thousands of copies of the mitochondrial genome dispersed throughout the mitochondrial network within a cell. In many metazoans, the mitochondrial genome has been reduced to just a few genes

C. Freyer · P. Clemente · A. Wredenberg (✉)
Division of Molecular Metabolism, Department of Medical Biochemistry and Biophysics, Karolinska Institute, Stockholm, Sweden
e-mail: Anna.Wredenberg@ki.se

J. Cruz-Reyes, M. Gray (eds.), *RNA Metabolism in Mitochondria*, Nucleic Acids and Molecular Biology 34, https://doi.org/10.1007/978-3-319-78190-7_2

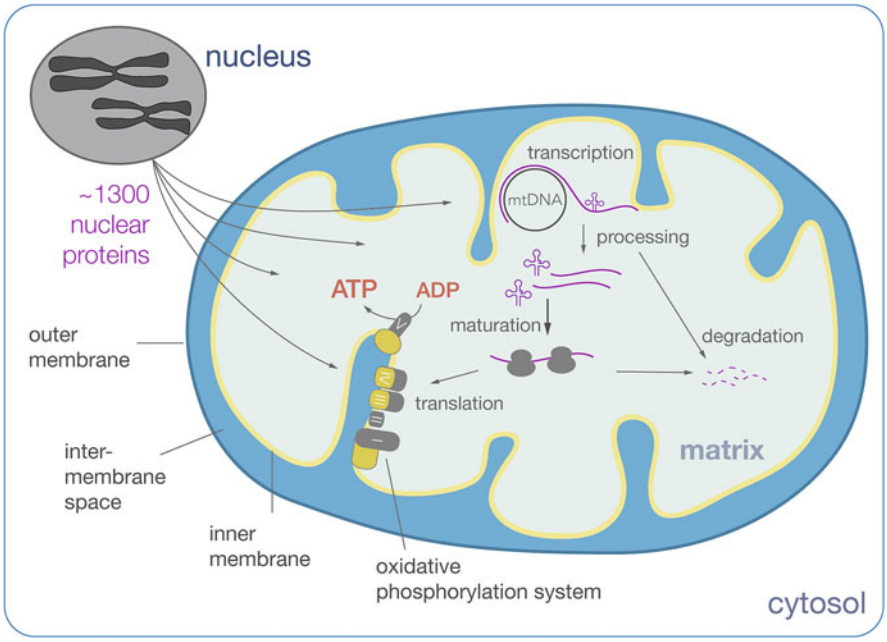

Fig. 2.1 Dual genetic control of mitochondrial function. The oxidative phosphorylation system (OXPHOS) is composed of subunits encoded by the nuclear as well as mitochondrial (mt) genomes. Thirteen subunits are encoded by mtDNA, transcribed and translated inside the mitochondrial matrix before being assembled in the inner mitochondrial membrane. Nucleus-encoded subunits are translated in the cytosol and imported into mitochondria. All factors for mitochondrial gene expression and translation are encoded in the nucleus and, together with around 1300 other proteins, actively imported into mitochondria

essential to support a respiratory chain in the mitochondrial membrane, as well as two structural RNAs for mitochondrial ribosomes. This cooperative arrangement requires a great deal of coordination between the nuclear and mitochondrial genomes, where over a thousand gene products need to be imported into mitochondria (Calvo et al. 2015) (Fig. 2.1). However, what governs the signals, both from the nucleus and from mitochondria in coordinating this nuclear-mitochondrial crosstalk, remains largely unknown. Similarly, some aspects of mitochondrial gene expression and its regulation are almost completely unknown. This review will discuss some recent advances.

Polyploidy varies between cell types and species but can range from just a few copies in single cellular organisms to over half a million copies of mitochondrial DNA (mtDNA) in a single mature oocyte. In most systems, the mitochondrial genome is a circular molecule (or circular mapping) within the mitochondrial matrix, ranging anywhere between 1 and 2000 kb in size, although linear genomes can be found in some plants, fungi and protists. These copies are usually identical within a cell and organism, a concept termed homoplasmy, while the presence of different

genotypes within an organelle is termed heteroplasmy. In rare cases, such as some plants and protists, different mtDNA molecules encode for different genes within the mitochondrial matrix (Kolesnikov and Gerasimov 2012). Heteroplasmy also plays an important role in human disease and ageing, and most organisms have developed mechanisms to ensure strict uniparental inheritance of their mitochondrial genome (Larsson 2010; Stewart and Chinnery 2015). In metazoans, the mitochondrial genome ranges between 11 and 28 kb in size, and for the purpose of this review, we will limit our discussion to findings made in humans, mice and *Drosophila melanogaster* (Dm).

2.1 Mitochondrial Function

Mitochondria have evolved to form highly dynamic networks within almost all eukaryotic cells, capable of fusing and budding to optimise and respond to the specific needs of each cell and even subcellular regions (Sheng and Cai 2012; Mishra and Chan 2014). Their outer membrane is permeable to most substrates and proteins smaller than 5 kDa, while the inner mitochondrial membrane (IMM) is obligatorily impermeable to almost all substrates, including electrons and protons. The latter character is essential for their function as energy-transducing membranes, where electrons from the citric acid cycle are transferred through a respiratory chain—a cascade of four protein complexes—to molecular oxygen, simultaneously pumping protons across the inner mitochondrial membrane. This mitochondrial respiration is tightly coupled to ATP synthesis (Kalckar 1939), and in 1961, Peter Mitchell formulated his chemiosmotic theory, suggesting that the generated electrochemical gradient is utilised by an ATP synthase upon proton re-entry into the mitochondrial matrix to synthesise ATP, from ADP and molecular phosphate (Mitchell 1961). The structures and molecular mechanisms of these five complexes, collectively called the oxidative phosphorylation (OXPHOS) system, have since been resolved (Tsukihara et al. 1995; Iwata et al. 1998; Zhu et al. 2016), with some notable exceptions, such as the exact mechanisms how the reduction of ubiquinone is coupled to the translocation of protons across the inner mitochondrial membrane at the NADH:ubiquinone oxidoreductase, or Complex I (Zhu et al. 2016).

By now, the endosymbiosis between the once-distinct cells has gone beyond just energy metabolism and involves a large number of diverse pathways, including glucose and lipid metabolism, steroid and haem synthesis, the production of heat and reactive oxygen species (ROS), cell signalling, apoptosis or calcium buffering (Chandel 2015). In total more than 1300 proteins are localised to mitochondria via a number of multisubunit complexes that regulate the import into the different mitochondrial compartments (Fig. 2.1) (Calvo et al. 2015; Wiedemann and Pfanner 2017). It is this central position in human metabolism that has placed mitochondria as a key factor in many human diseases, and mitochondrial dysfunction can be seen in a wide variety of clinical presentations as well as in the natural ageing process

(Larsson 2010; Nunnari and Suomalainen 2012; Bratic and Larsson 2013; Levin and Mishmar 2015; Kauppila et al. 2017).

2.2 The Mitochondrial Genome

In humans, as well as most other metazoans, the mitochondrial genome is a small, circular, double-stranded, multicopy genome, dispersed throughout the mitochondrial network (Fig. 2.2). In reference to their bacterial ancestry, mtDNA is compacted into supercoiled structures called nucleoids (Satoh and Kuroiwa 1991) of roughly 100 nm in size (Kukat et al. 2011; Brown et al. 2011), and studies in a variety of species have since demonstrated that the high-mobility group (HMG) box protein, mitochondrial transcription factor A (TFAM), is responsible for this packaging (Parisi et al. 1993; Larsson et al. 1998; Alam et al. 2003; Matsushima et al. 2003; Kukat et al. 2015; recently reviewed in Kukat and Larsson 2013). TFAM levels closely resemble mtDNA copy number, and manipulating TFAM levels in vivo can alter mtDNA copy number (Larsson et al. 1998; Ekstrand et al. 2004; Wai et al. 2010; Freyer et al. 2010; Bratic et al. 2015). In most animals, mtDNA has a strict uniparental inheritance, with transmission of the mitochondrial genome through the female germ line (Ladoukakis and Zouros 2017). Mature mammalian

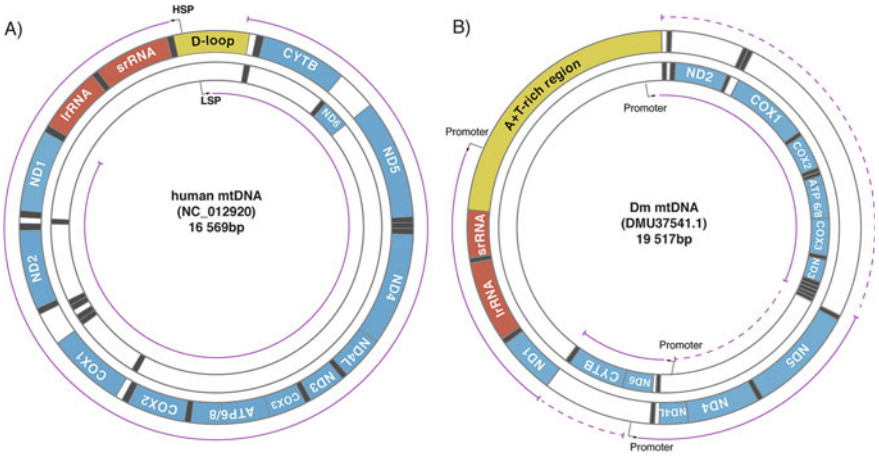

Fig. 2.2 Gene structure of the (**a**) human and (**b**) *Drosophila melanogaster* (Dm) mitochondrial genomes. Control regions, mRNAs, tRNAs, and rRNAs are shown in yellow, blue, grey, and red, respectively. In human mitochondria, transcription is initiated on both strands at the heavy- and light-strand promoters, HSP and LSP, respectively. Transcription then proceeds around almost the entire genome before terminating within the regulatory region (D-loop) or the antisense region of the rRNA transcripts. Transcription in Dm is less well understood, but several transcription units (solid lines), promoters and termination sites have been proposed. The existence of promoters outside of the control region (A + T-rich region) is not clear, but this model would suggest the absence of antisense RNA (dashed lines)

oocytes contain a surprisingly high number of mtDNA molecules, and the transmission of paternal mtDNA is prohibited by a number of mechanisms (Shoubridge and Wai 2007; Stewart et al. 2008; Sato and Sato 2013; Stewart and Larsson 2014). For instance, factors regulating mtDNA copy number are significantly downregulated during spermatogenesis (el Meziane et al. 1989; Larsson et al. 1997; Rantanen et al. 2001; Amaral et al. 2007), while paternal mitochondria are ubiquitinated and actively degraded in the zygote (Sutovsky et al. 1999). In Dm paternal mtDNA is already removed during spermatogenesis to ensure a strict maternal inheritance of mtDNA (DeLuca and O'Farrell 2012), while in *Drosophila simulans*, paternal contribution seems to be an integral part of its inheritance (Wolff et al. 2013). Although the reasons for this strict evolutionary behaviour is unclear, heteroplasmy can have major implications in human disease and ageing (Ma et al. 2014; Stewart and Chinnery 2015).

2.3 Replication

The two strands of human mtDNA can be separated by buoyant density ultracentrifugation, due to their differential cytosine/guanosine content, historically designating the two strands as either the heavy (H) or light (L) strands (Battey and Clayton 1978). Gene content and order vary among species, but in most bilaterian animals, genes are distributed on both strands, encoding 13 essential subunits of the OXPHOS system together with 2 rRNAs and 22 tRNAs necessary for mitochondrial translation. Gene structure in most of these mitochondrial genomes is highly compact, and human mtDNA has no introns and only one major non-coding region that contains promoter regions as well as the origin of H-strand replication (O_H) (Gustafsson et al. 2016). The remaining genome is highly compact, with no or only little separation between genes. Open reading frames can even be overlapping and expressed from the same bicistronic transcript. The H-strand encodes the majority of transcripts, specifying 10 mRNAs, 14 tRNAs and both ribosomal subunits, while the L-strand encodes 1 additional mRNA and 8 tRNAs (Fig. 2.2). In contrast, gene organisation in *D. melanogaster* is slightly different with a more even gene content distribution between the two strands (Lewis et al. 1995).

Replication of the mitochondrial genome is catalysed by a single replication complex, with many factors related to bacteriophage homologs. The human mtDNA polymerase holoenzyme consists of mtDNA polymerase γ (POLγ), comprising a catalytic subunit (POLγA) and two accessory subunits (POLγB), while in *D. melanogaster*, the holoenzyme is composed of one catalytic and one accessory subunit (Kaguni 2004). The basic mitochondrial replisome requires the mitochondrial helicase (Twinkle) and the single-stranded binding (SSB) proteins for complete mtDNA replication [recently summarised in Young and Copeland (2016) and Gustafsson et al. (2016)], and although POLγ, Twinkle and SSB are sufficient to replicate mtDNA in vitro, additional factors seem to be involved for faithful replication in vivo. For instance, the 5′-3′ exonuclease MGME1 and DNA ligase III have

been shown to be required to join the ends of the newly synthesised mtDNA molecules to generate its circular form (Lakshmipathy and Campbell 1999; Puebla-Osorio et al. 2006; Bratic et al. 2015; Macao et al. 2015; Uhler et al. 2016). An additional DNA polymerase, referred to as PrimPol (primase-polymerase), has been shown to be able to resolve stalling at replication blocks and reinitiate DNA synthesis from aborted templates but is unable to rescue the loss of POLγ (Torregrosa-Muñumer et al. 2017).

2.4 Transcription

Mitochondrial transcription is initiated from two designated promoter regions localised in the control region in both vertebrates and most likely also in invertebrates (Garesse and Kaguni 2005), generating long primary polycistronic transcripts covering almost the entire genome (Montoya et al. 1982; Chang and Clayton 1984; Bogenhagen et al. 1984). The circular nature of most mitochondrial genomes inevitably poses the possibility of collisions during simultaneous transcription, requiring tight regulation. H-strand transcription has been proposed to be initiated from two promoters ($HSP_{1,2}$), although there is no consensus regarding the existence of HSP_2 (Gustafsson et al. 2016). However, it is clear that two overlapping transcripts with a 50-fold difference in steady-state levels are generated from HSP, with the preferentially expressed, shorter transcript being initiated upstream of the $tRNA^{Phe}$ gene and containing both ribosomal subunit rRNAs, of 12S and 16S, before terminating within the $tRNA^{Leu(UUR)}$ gene. The second polycistronic transcript is proposed to initiate from within the $tRNA^{Phe}$ gene and to include the remaining H-strand transcripts (Fig. 2.2). Termination of the longer transcript coincides with specific termination sequence blocks, but the exact mechanism is not yet known (Freyer et al. 2010; Jemt et al. 2015). In contrast, termination at $tRNA^{Leu(UUR)}$ has been proposed to be due to binding of the mitochondrial termination factor, MTERF1, but other results have suggested that MTERF1 might prevent L-strand expression of antisense rRNA sequences and direct replication fork pausing (Kruse et al. 1989; Asin-Cayuela et al. 2005; Hyvärinen et al. 2007; Yakubovskaya et al. 2010; Terzioglu et al. 2013; Shi et al. 2016).

Transcription of L-strand transcripts is initiated at the light-strand promoter (LSP) and extends throughout the genome before terminating at or close to the MTERF1-binding site. Additionally, LSP transcription forms the primer required for mtDNA replication by terminating ~100 nt downstream of LSP at conserved sequence blocks, responsible for the transition from RNA to DNA (Xu and Clayton 1995, 1996; Pham et al. 2006; Wanrooij et al. 2010, 2012). The mechanism has recently been reviewed elsewhere (Gustafsson et al. 2016). Mitochondrial transcription can be affected by a number of environmental stimuli, as well as by manipulating the expression of a range of different factors, either stimulating or inhibiting de novo mitochondrial transcription. In *Drosophila*, several polycistronic transcription units have been reported, covering either strand. Two members of the MTERF

transcription termination family, mTTF and mTERF5, have been shown to interact with two sequence elements at the boundaries of these transcription units regulating transcription termination with opposing effects (Roberti et al. 2006; Bruni et al. 2012). Transcription in *Drosophila* either initiates from the A + T-rich region, similar to the situation in humans, or from several sites within the genome (Roberti et al. 2006). Currently it is not clear which model is correct, but disrupting the factors involved in regulation might provide better evidence.

Despite its α-proteobacterial heritage, the metazoan mitochondrial transcription machinery, just as its replication counterpart, is much closer related to the T7 bacteriophage system (Falkenberg et al. 2007). Besides the mitochondrial RNA polymerase (POLRMT) (Ringel et al. 2011) and TFAM (Parisi and Clayton 1991; Dairaghi et al. 1995; Malarkey et al. 2012), the mitochondrial transcription factor 2 (TFB2M) is essential for mitochondrial transcription initiation (Falkenberg et al. 2002; Sologub et al. 2009; Litonin et al. 2010), and deletion of these factors is lethal in early development (Larsson et al. 1998; Adán et al. 2008; Kühl et al. 2014). The current model for transcription suggests that TFAM binds and bends mtDNA at the promoter regions (Kukat and Larsson 2013), before recruiting POLRMT and allowing TFB2M to melt the promoter region. Whether TFAM and TFB2M jointly are required is not yet fully resolved (Ramachandran et al. 2017; Posse and Gustafsson 2017), but once transcription commences, TFB2M is released, and the mitochondrial transcription elongation factor, TEFM, is recruited to the remaining transcription complex, where it aids the transcription machinery to navigate through structurally difficult regions. TEFM might even regulate the switch between transcription and replication at the CSBs (Minczuk et al. 2011; Posse et al. 2015; Hillen et al. 2017). Finally, the mitochondrial ribosomal protein L7/L12 (MRPL12) has been proposed to interact with POLRMT to stimulate transcription and allow full transcription of the polycistronic transcripts (Nouws et al. 2016). The *Drosophila* mitochondrial genome does not contain such a defined regulatory region, but instead has a long A + T-rich region, where no structural features have been reported.

2.5 Processing and Maturation

The current model of the life cycle of a metazoan mitochondrial RNA is that the polycistronic transcripts are processed into their individual units by processing enzymes that recognise the junctions, which are often delineated by tRNAs (Ojala et al. 1981) (Fig. 2.3). These transcripts are then stabilised by an RNA-binding complex, allowing for further maturations, before being targeted to mitochondrial ribosomes for translation, after which the mitochondrial degradosome degrades unwanted transcripts. The basic machinery and many factors involved in this process have been identified, but how these factors are connected in a spatiotemporal manner is not clear. Nor is this simplified view consistent with some observations made experimentally, requiring further investigation. For instance, although the molecular

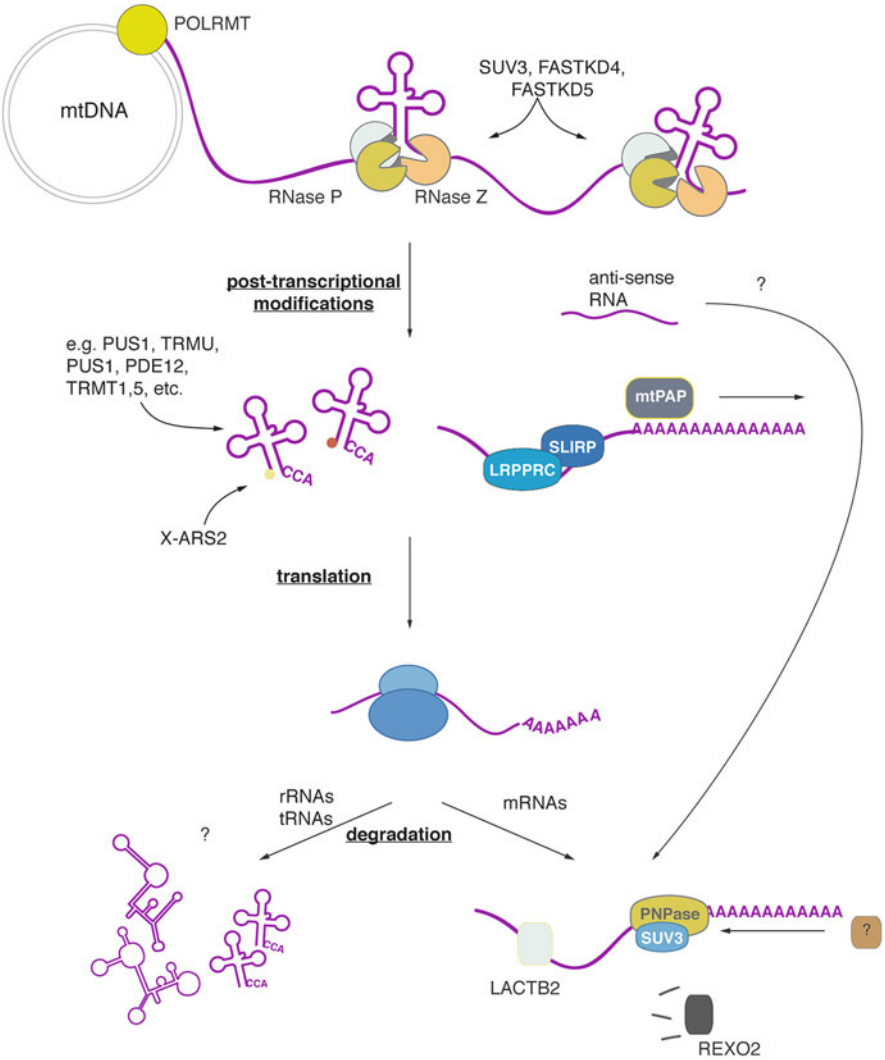

Fig. 2.3 The life cycle of mitochondrial RNA. The mitochondrial RNA polymerase (POLRMT) generates long, polycistronic transcripts, which are processed by RNase P and RNase Z, possibly with the help of protein factors SUV3, FASTKD4 and FASTKD5. Coding transcripts are stabilized by the LRPPRC/SLIRP complex for modifications, such as polyadenylation by mitochondrial poly (A) polymerase (MTPAP). Mature mRNAs are translated close to the inner mitochondrial membrane before being degraded by a combination of polynucleotide phosphorylase (PNPase), SUV3, LACTB2 and RNASET2. How antisense RNAs are targeted for degradation, and whether deadenylases regulate degradation, is not known. Neither is it clear how tRNAs and rRNAs are degraded

details of transcription initiation and processing of the polycistronic transcripts are becoming more clear, basic principles of how these transcripts are extracted from the nucleoid, where and when they are processed and modified and how translation is initiated remain largely unknown.

Some aspects of how these processes could be organised are slowly emerging. Perhaps not surprisingly, newly synthesised RNAs localise in close proximity to nucleoids (Iborra et al. 2004), and it was later shown that a number of factors involved in mtRNA metabolism and translation can also be found at these sites, which gave rise to the concept of mitochondrial RNA granules (MRGs) (Borowski et al. 2013; Antonicka et al. 2013; Jourdain et al. 2013, 2015; Lee et al. 2013; Bogenhagen et al. 2014; Wilson et al. 2014; Antonicka and Shoubridge 2015; Tu and Barrientos 2015; Zaganelli et al. 2017). RNA granules are ribonucleoprotein particles, present in a variety of systems, which are often visible by light microscopy. In the nucleus, cytosol and chloroplast, RNA granules have been reported to form in response to stress or for RNA modifications and rRNA transcription. The function of MRGs is less clear, as factors involved in processing, modification, mitoribosome assembly, translation and RNA degradation have all been localised to MRGs [recently summarised in Jourdain et al. (2016)]. In the absence of RNA, however, these structures disappear, suggesting that their organisation might be less strict. Some form of sub-MRG organisation has also been suggested, where so-called D-foci form the site of RNA degradation (Borowski et al. 2013).

2.5.1 The tRNA Punctuation Model

In the canonical mitochondrial tRNA punctuation model, mRNAs on the long polycistronic transcripts are flanked by tRNAs, which are recognised and released by endonucleolytic cleavage (Ojala et al. 1980, 1981). This step is essential for mRNA maturation, and the two mitochondrial endonucleases, RNase P and RNase Z, consecutively process the polycistronic transcript, releasing short immature transcripts (Sanchez et al. 2011; Rossmanith 2012; Reinhard et al. 2017). Unlike other RNase P complexes, the mitochondrial RNase P-like complex does not contain an RNA component but is made up of three mitochondrial RNase P peptides, MRPP1–3 (see also Chap. 3). In vitro studies demonstrated that MRPP1 and 2 form a complex at the 5′ tRNA junction to perform an N^1-methylation of m^1A9 or m^1G9 of the tRNA stem (Helm et al. 1998, 1999; Vilardo et al. 2012). Besides its involvement in RNA processing, MRPP2 also has dehydrogenase function and has been shown to be involved in amino acid catabolism and lipid metabolism (Shafqat et al. 2003; Yang et al. 2014). MRPP3 is dependent on these modifications to perform nucleolytic cleavage at the 5′ end of tRNA. Disrupting MRPP1, 2 or 3 in

Drosophila (Sen et al. 2016) or MRPPR3 in the mouse (Rackham et al. 2016) leads to lethality and the accumulation of processing intermediates. Mutations in RNase P have also been associated with mitochondrial diseases in humans (Vilardo and Rossmanith 2015; Falk et al. 2016; Metodiev et al. 2016).

Mitochondrial RNase Z, on the other hand, is composed of the metallo-beta-lactamase, ELAC2, which mediates 3′ processing of tRNAs (Brzezniak et al. 2011; Sanchez et al. 2011; Rossmanith 2012; Xie and Dubrovsky 2015). Further, studies made in flies (Dubrovsky et al. 2004; Xie and Dubrovsky 2015; Andreenkov et al. 2016) and mice (Rackham et al. 2016) supported the notion that 5′ processing preceeds 3′ cleavage. The MRPP1/2 complex has recently been shown to also promote 3′ cleavage by RNase Z, by presenting the released tRNA to the CCA-adding enzyme (Reinhard et al. 2017). MRPP1 and 2 therefore function as a platform, where mitochondrial tRNAs are processed and matured, and it will be interesting to see whether other tRNA modifications are also dependent on MRPP1/2.

Not all mRNAs are flanked by tRNAs, and thus the canonical processing of polycistronic transcripts requires modification. Mitochondrial systems with severely reduced tRNA content, such as trypanosome or coral mitochondrial genomes, which contain no or only single tRNAs, might be important models to help identify the mechanisms involved in the processing of noncanonical sites. In light of this requirement, additional components important for correct processing of mitochondrial primary transcripts have recently been identified. For instance, the pentatricopeptide domain-containing protein, PTCD1, is thought to regulate $tRNA^{Leu}$ levels in mitochondria and has also been implicated in the processing of primary transcripts, but more work is required to understand how PTCD1 is involved in this process (Sanchez et al. 2011). Similarly, loss of PTCD2 in the mouse heart showed defective 5′ cleavage of mitochondrial apocytochrome *b* transcript, suggesting a role in processing (Xu et al. 2008). The mitochondrial helicase SUV3 is primarily implicated to be part of the mitochondrial degradosome (see below), but depletion in *D. melanogaster* (Clemente et al. 2015), yeast (Zhu et al. 1989; Stepien et al. 1992; Hoffmann et al. 2008) or human cell lines (Szczesny et al. 2010) resulted in the accumulation of unprocessed polycistronic transcripts and a failure to accumulate mature tRNAs, strongly indicating that helicase activity is required for efficient processing.

2.6 Maturation and Modification

2.6.1 tRNAs

Base and sugar modifications are an essential feature of RNAs, and ~7% of all tRNA positions are modified. For the vast majority of these modifications, the responsible enzymes have not yet been identified (Suzuki and Suzuki 2014; Van Haute et al. 2015). Pseudouridylation is the most common modification of

non-coding RNAs, with modifications reported in all mt-tRNAs as well as 16S rRNA. However, the function, the effect on structure or the enzymes responsible are largely unknown, with the exception of PUS1, which pseudouridylates positions 27 and 28 of tRNAs (Patton et al. 2005; Fernández-Vizarra et al. 2007). Additional putative pseudouridylate synthases are RPUSD3 and 4 and TRUB2 (Arroyo et al. 2016; Zaganelli et al. 2017). Identifying their functions will be important to understand tRNA modifications. In contrast to the cytosol, mitochondria express only one $tRNA^{Met}$, both for elongation and translation initiation. Both forms require a 5-formyl modification at the wobble base C34 (Takemoto et al. 2009; Bilbille et al. 2011), which is performed by the stepwise actions of the $tRNA^{Met}$-(m5C34)-methyltransferase NSUN3 (Van Haute et al. 2016; Haag et al. 2016) and the Fe^{2+}/α-ketoglutarate-dependent deoxygenase, ALKBH1 (Haag et al. 2016). Whether α-ketoglutarate, a key component of the Krebs cycle, is therefore capable of regulating mitochondrial translation is yet unclear. Finally, the mitochondrial nucleotidyltransferase 1 (TRNT1) is responsible for 3′-CCA addition by polymerising CTP and ATP onto the 3′ discriminator base of tRNAs (Nagaike et al. 2001; Augustin et al. 2003), an essential modification during tRNA maturation.

2.6.2 rRNA

Similarly, several methylation modifications on both the small (Seidel-Rogol et al. 2003; Metodiev et al. 2009, 2014) and large (Lee et al. 2013; Lee and Bogenhagen 2014; Rorbach et al. 2014; Bar-Yaacov et al. 2016) mitoribosomal rRNAs have been identified, which are essential for rRNA stability and monosome assembly.

2.6.3 mRNA

Base modifications on mRNAs, such as methylations or pseudouridylations, are rare, although a recent report demonstrated a m^1A1374 modification in the ND5 mRNA, expressing a subunit of human complex I (Safra et al. 2017). On the other hand, mitochondrial mRNAs, but also tRNAs and rRNAs, can undergo a process termed RNA editing, where the coding sequence of the mature RNA differs from the initial DNA sequence. These edits can range from single nucleotide exchanges to entire stretches that are deleted and replaced by a different sequence. These latter forms are widespread in plant and trypanosome mitochondria, but not reported in humans. In mammals single nucleotide modifications on mtRNAs have been reported, such as cytidine-to-uridine or adenosine-to-inosine deaminations and have been discussed elsewhere (Chateigner-Boutin and Small 2011).

2.7 Polyadenylation of Mitochondrial mRNAs

One of the most extensive post-transcriptional RNA processing events is the polyadenylation of mature transcripts, where non-templated adenosine residues are added to the 3′ end of RNAs. Despite its common prokaryotic origin, the addition of the poly(A) tail has different effects in the post-transcriptional regulation of transcripts across the different kingdoms and different organelles. In bacteria and plant organelles, polyadenylation is considered to signal degradation of the mRNAs (Schuster and Stern 2009). In these cases, the 3′ termini of mature mRNAs are protected from exonucleolytic attack by a structural terminator stem loop. Polyadenylation by poly(A) polymerases such as *pcnB* in *E. coli* or ASG1 in plants relaxes these structures and provides access for exonucleases (Hirayama et al. 2013; Hui et al. 2014). In the nucleus and cytosol of eukaryotes, poly(A) tails are present in the majority of mRNAs and play an important role in their stability, export from the nucleus and translation (Weill et al. 2012). Yeast mt-mRNAs, on the other hand, have developed a different mechanism and are not polyadenylated (Groot et al. 1974) but instead contain an AU-rich (*Saccharomyces cerevisiae*) or C-rich (*Schizosaccharomyces pombe*) dodecamer sequence at the 3′ end, which promotes transcript stability (Butow et al. 1989). RNA-binding proteins are proposed to interact with these sequences, preventing degradation by the mitochondrial RNA degradosomes.

In human mitochondria, however, the role of polyadenylation is less clear. More than 45 years ago, polynucleotide adenylyltransferase activity was reported in rat liver mitochondrial extracts (Jacob and Schindler 1972), capable of catalysing the addition of AMP to 3′ termini of RNA from ATP in a template-independent manner. Since then, the gene responsible has been isolated, and despite its common origin, the poly(A) polymerase found in protist and animal mitochondria differs from its bacterial counterpart. Mitochondrial poly(A) polymerases (MTPAP, also known as PAPD1 or TUTase1) belong to the group of noncanonical DNA type b polymerases, capable of polyadenylation as well as polyuridylation, and have been characterised in a number of species, including humans (Tomecki et al. 2004) and flies (Bratic et al. 2016). In most bilaterian mitochondria, MTPAP adds 40–50 adenylate residues to the 3′ termini of almost all mitochondrial transcripts, although polyuridylation and polycytidylations have been reported, respectively, in mitochondria of trypanosomes and *Chlamydomonas* (Aphasizhev and Aphasizheva 2008; Zimmer et al. 2009; Salinas-Giegé et al. 2017). The crystal structures from humans and chickens revealed that MTPAP functions as a homodimer without the need for additional protein cofactors (Bai et al. 2011; Lapkouski and Hallberg 2015), although recent data suggest that MTPAP activity can be stimulated by other factors such as SUV3 (Wang et al. 2014) and the leucine-rich PPR motif-containing protein, LRPPRC (Wilson et al. 2014) (see below). With the exception of ND6 mRNA, which has a long 3′ untranslated region (UTR), all mature human and murine mt-mRNAs are polyadenylated (Temperley et al. 2010; Ruzzenente et al. 2012). Seven out of the 13 open reading frames do not encode a functional stop codon, and thus

polyadenylation is required for completion of a UAA stop codon (Temperley et al. 2010). The reason for the absence of polyadenylation of the ND6 transcript in humans and rodents is not entirely clear, and ND6 mRNA is indeed polyadenylated in other species, including birds (Sun et al. 2017) and flies (Stewart and Beckenbach 2009; Bratic et al. 2011). In these cases, gene rearrangements have restructured the mitochondrial genomes, potentially introducing defined processing signals at the 3′ termini and promoting polyadenylation. Besides ND6 mRNA, the transcript of ND5 has a long 3′ UTR with a short poly(A) tail in humans, potentially removing the necessity of long poly(A) tails (Temperley et al. 2010; Mercer et al. 2011). The two mitochondrial rRNAs contain short poly(A) tails in humans, while in flies, 16S rRNA has a long poly(A) tail of around 40 adenosines (Bratic et al. 2011, 2016; Clemente et al. 2015). Whether this elongated 3′ terminus of 16S RNA has any structural consequences has not yet been explored.

As mentioned above, polyadenylation can serve as a signal for degradation in bacteria, as well as plant organelles. However, in these cases, only a small proportion of mature transcripts are actually polyadenylated, reflecting their status in RNA turnover. In contrast, in most bilaterian mitochondria, almost all transcripts are polyadenylated, suggesting a different role for the poly(A) tail, for example, in promoting translation or transcript stability. Mitochondrial transcripts are not capped and in general do not require circularisation for translation, and thus a direct role in translation initiation has not been attributed to the poly(A) tail. In contrast, numerous studies have addressed its role in transcript stability, and inactivation of MTPAP leads to a shortening of the mitochondrial poly(A) tails with opposite effects on the stability of mitochondrial transcripts. These results led to the suggestion that mitochondrial polyadenylation has both stabilising and destabilising roles. In general, the inactivation of MTPAP by RNAi in human cell lines leads to increased steady-state levels of some ND mRNAs, decreased steady-state levels of COX transcripts and unchanged levels of CytB mRNA (Tomecki et al. 2004; Nagaike et al. 2005; Piechota et al. 2006; Nagao et al. 2008). These results seemed to be corroborated by the identification of a patient with severely compromised MTPAP activity, leading to spastic ataxia (Crosby et al. 2010; Wilson et al. 2014). The mitochondrial transcripts analysed in the forgoing studies all retained short adenylate stretches at their 3′ ends and, together with the low processivity of MTPAP observed in vitro, led to the suggestion that an additional putative enzyme is required to add these short oligoadenylate stretches prior to MTPAP function and normal polyadenylation. However, using an MTPAP knockout *D. melanogaster* model, we failed to observe such oligoadenylation activity on mRNAs, questioning the presence of such an enzyme in mitochondria (Bratic et al. 2016). These DmMTPAP KO flies presented with a complete absence of adenylation on mt-mRNAs, and transcripts were trimmed up to 20 nucleotides (nt) at their 3′ ends. The only transcripts that retained some degree of polyadenylation were the two ribosomal RNAs, 12S and 16S, which either retained an unchanged tail length of ~5 nt or had reduced polyadenylation relative to the usual ~40 nt tail, respectively (Bratic et al. 2016). MTPAP KO flies are larvae lethal with a severe mitochondrial dysfunction, and with the exception of transcripts for DmCOX1 and DmCYTB, which were unchanged or reduced,

respectively, all mt-mRNA steady-state levels were increased despite their compromised 3′ integrity (Bratic et al. 2016). This observation, together with results obtained in human cell lines (Wydro et al. 2010), suggests that polyadenylation by MTPAP is not required to protect mitochondrial transcripts from degradation. Furthermore, de novo labelling experiments also excluded the possibility that polyadenylation was required for translation and that despite the truncated 3′ ends, mitochondrial translation could still occur (Bratic et al. 2016).

2.8 Mitochondrial RNA Stability

RNA turnover is determined by the rate of de novo transcription and degradation. Neither of these processes is constant, and different RNA species have different half-lives, depending on the type of RNA and the metabolic demands. Thus, both transcription and degradation are highly regulated by protein complexes, such as exosomes and degradosomes (Houseley and Tollervey 2009). In mitochondria, it is reasonable to believe that RNA turnover is an important form of regulating gene expression and transcripts can be rapidly removed or increased under the correct physiological stimuli. Besides leading to mature mRNA, tRNA and rRNA transcripts, mitochondrial transcription also results in large amounts of non-coding, misfolded and aborted transcripts, as well as in processing by-products such as mirror or antisense RNAs. However, despite deriving from the same polycistronic transcript, the latter RNA species are rarely detected under normal physiological conditions (Aloni and Attardi 1971; Wang et al. 2009; Szczesny et al. 2010; Borowski et al. 2013). Thus, the different transcripts generated during transcription have different half-lives, but what governs their half-life is not entirely clear. One factor that is described as promoting RNA stability in mitochondria is the leucine-rich PPR motif-containing protein, LRPPRC. Pentatricopeptide repeat (PPR) proteins form a large family of modular RNA-binding proteins, which are involved in various aspects of gene expression. In plants, over 400 different PPR domain-containing proteins regulate RNA synthesis, processing and degradation in the nucleus and organelles (Barkan and Small 2014), while in humans, only 7 members have been identified, all localising to mitochondria (Rackham and Filipovska 2012; Lightowlers and Chrzanowska-Lightowlers 2013). Besides LRPPRC, these PPR proteins include PTCD 1–3, POLRMT (the mitochondrial RNA polymerase), the RNase P subunit MRPP3 as well as the mitochondrial ribosomal protein, MRPS27. LRPPRC has 35 PPR domains and was initially identified as interacting with and stabilising polyadenylated mitochondrial transcripts (Mili and Piñol-Roma 2003). Since then, it has become clear that LRPPRC forms a stable complex with the stem-loop interacting RNA-binding protein, SLIRP (Sasarman et al. 2010; Bratic et al. 2011; Ruzzenente et al. 2012; Chujo et al. 2012; Wilson et al. 2014; Lagouge et al. 2015; Spåhr et al. 2016; Siira et al. 2017), and deletion of LRPPRC leads to a severe

mitochondrial dysfunction and loss of mitochondrial mRNAs (Gohil et al. 2010; Bratic et al. 2011; Xu et al. 2012; Ruzzenente et al. 2012; Harmel et al. 2013; Mourier et al. 2014). Additionally, disruption of LRPPRC in mice (Ruzzenente et al. 2012) and flies (Bratic et al. 2011) led to a loss of the poly(A) tail of mitochondrial transcripts, and it was later shown that LRPPRC is required for polyadenylation, as well as stimulating MTPAP activity in vitro (Chujo et al. 2012; Wilson et al. 2014; Siira et al. 2017).

However, although it is clear that LRPPRC is able to stabilise mitochondrial mRNAs and plays a role in their maturation, it is not clear whether LRPPRC also regulates RNA degradation by distinguishing between coding transcripts and RNAs destined for degradation. In vitro work with recombinant LRPPRC, as well as RNA precipitation experiments, does not suggest that LRPPRC is able to distinguish between different RNA species (Spåhr et al. 2016; Siira et al. 2017). Thus, it is more likely that the decision to keep or destroy a transcript is already made prior to LRPPRC binding or made by other interacting proteins. An interesting candidate would have been SLIRP, with which LRPPRC forms a stable complex (Sasarman et al. 2010; Ruzzenente et al. 2012; Chujo et al. 2012; Lagouge et al. 2015; Spåhr et al. 2016), but the knockout of SLIRP in a mouse model did not increase antisense RNAs but rather suggested that this protein is responsible for targeting mature transcripts for translation (Lagouge et al. 2015).

Drosophila contain two homologs of both LRPPRC and SLIRP (Sterky et al. 2010; Bratic et al. 2011). DmLRPPRC1, also known as the bicoid stability factor (BSF), shares the same phenotypes as its mammalian homolog when disrupted (Bratic et al. 2011), and recent experiments in Dm Schneider S2 cells demonstrated that DmLRPPRC1 is degraded by the mitochondrial matrix protease ClpXP, potentially linking protein degradation to gene expression (Matsushima et al. 2017). In contrast to its paralog, though, silencing of DmLRPPRC2 does not affect mRNA stability or polyadenylation but rather seems to affect the coordination of translation (Baggio et al. 2014).

A number of RNA-binding proteins belonging to the family of Fas-activated serine/threonine kinases (FASTK, FASTKD1–5) have recently been localised to human mitochondria, and although their functions have not yet been fully resolved, they all seem to be involved in mtRNA metabolism (Jourdain et al. 2017). Interestingly, in contrast to LRPPRC, which does not have any sequence specificity, FASTKs do seem to act differentially on transcripts. FASTK has been reported to stabilise ND6 mRNA, by interacting with its 3′ end (Jourdain et al. 2015), while FASTKD1 promotes the degradation of the ND3 transcript (Boehm et al. 2017). FASKD2 and 3 are required for efficient mitochondrial translation (Popow et al. 2015; Boehm et al. 2017), and FASKD4 and 5 are reported to be involved in the processing of noncanonical transcript junctions (Wolf and Mootha 2014; Antonicka and Shoubridge 2015). *Drosophila* has only three orthologs, which all localise to mitochondria, but their functions have not yet been described.

2.8.1 Controlled mtRNA Degradation

Degradation of mRNAs is a major mechanism for regulating gene expression. In bacteria, transcripts have a rapid turnover, with half-lives of only 3–8 min for the vast majority (Bernstein et al. 2002), while in mouse fibroblasts, RNAs can be retained for over 20 h (Schwanhäusser et al. 2013). Cytosolic degradation is highly organised and occurs in designated structures, termed processing bodies (P-bodies), involved in mRNA decay, decapping or nonsense-mediated decay. Mitochondrial RNA decay has been proposed to occur in MRG substructures, termed D-foci (Borowski et al. 2013), but their organisation and regulation are much less well understood. In human mitochondria, mRNA half-lives have been calculated to range between 1 and 3 h in cell lines, depending on transcript and physiological conditions, but what controls degradation is not clear (Nagao et al. 2008; Borowski and Szczesny 2014). Two factors have been identified to form the minimal mitochondrial degradosome, consisting of the ATP-dependent RNA helicase SUV3 and polynucleotide phosphorylase (PNPase). SUV3 belongs to a highly conserved Ski2 family of DExH-box RNA helicases, with orthologs found in eukaryotes to *Rhodobacter* (Dmochowska et al. 1999), and was initially identified in yeast as a suppressor of the var1 (SUV3) phenotype (Butow et al. 1989). Var1 encodes a mitochondrial ribosomal subunit, and SUV3 was suggested to rescue the subsequent impaired translation by altering the levels of mitochondrial transcripts (Butow et al. 1989; Conrad-Webb et al. 1990). In yeast, SUV3 has been shown to work with the RNase II-like exonuclease, DSS1 (Stepien et al. 1992; Dziembowski et al. 2003). Higher eukaryotes have lost DSS1, and instead, SUV3 forms a complex with PNPase in order to degrade RNA in a 3′ to 5′ direction (Wang et al. 2009; Szczesny et al. 2010; Borowski et al. 2013). The ubiquitous disruption of SUV3 in mice is embryonic lethal, and heterozygous knockout animals have a reduced lifespan, increased tumour formation and reduced mtDNA levels with increased mtDNA mutation load (Chen et al. 2013). However, the effects on mitochondrial RNA turnover were not investigated in this study. The disruption of SUV3 in flies (Clemente et al. 2015) showed considerable agreement with studies performed in human cell lines, where silencing of the human *SUV3* gene product (SUPV3L1) or the expression of a dominant-negative variant in human cells resulted in increased mRNA steady-state levels, as well as the accumulation of mRNA decay intermediates, processing by-products and antisense RNAs (Szczesny et al. 2010). Additionally, both models demonstrated a loss of mature mt-tRNAs, due to the lack of correct processing of the primary polycistronic transcripts. Interestingly, disruption of SUV3 had opposing effects on the poly(A) tail, with shortening in flies (Clemente et al. 2015) but lengthening in human cells (Szczesny et al. 2010).

In plants PNPase is involved in the processing and regulation of polyadenylation-dependent mtRNA decay (Schuster and Stern 2009; Zimmer et al. 2009; Stoll et al. 2014) and has been shown to both degrade and extend 3′ tails in vitro (Nagaike et al. 2005). Despite its proposed interaction with the mitochondrial matrix protein SUV3, PNPase has been localised to the

intermembrane space (Chen et al. 2006), where it has been suggested to import RNAs, including 5S rRNA (Wang et al. 2010). However, this conclusion has been questioned as neither the RNase P complex (Lin et al. 2012; Rossmanith 2012; Karasik et al. 2016) nor the mammalian mitochondrial ribosome requires cytosolic non-coding RNAs (Amunts et al. 2015; Rorbach et al. 2016), and no clear evidence of cytoplasmic RNAs inside mitochondria have been demonstrated for human mitochondria. However, PNPase has also been localised to the D-foci substructures, suggesting at least a dual localisation for PNPase (Borowski et al. 2013). Loss of DmPNPase or DmSUV3 results in increased mt-mRNA levels and larvae lethality but unchanged mt-rRNA levels in *D. melanogaster* models (Clemente et al. 2015). While the overexpression of either DmSUV3 or DmPNPase has only mild effects on fly viability and mt-mRNA levels, double overexpression of both components of the degradosome leads to a severe mt-mRNA depletion, mild reduction of mt-rRNAs and larvae lethality (Pajak et al., unpublished data). These results further suggest a role for a SUV3-PNPase complex in the decay of mitochondrial RNAs, but what regulates this degradosome and how it distinguishes between the different classes of mitochondrial RNAs remain unclear. Recently, a study in human cell lines proposed that mtRNAs are actually targeted for degradation to the intermembrane space, where the ribonuclease RNASET2 rather than PNPase is responsible for mtRNA decay (Liu et al. 2017). How the RNAs are transported across the IMM is not clear, and further work is required. Additionally, the metallo-lactamase, LACTB2, has also been proposed to act as an endonuclease in mitochondria, and while its disruption led to only a mild increase in mtRNAs, in vitro experiments suggested a role in RNA degradation (Levy et al. 2016).

Several additional putative ribonucleases have been reported to act in mitochondria, but their roles are not fully understood. For instance, the mitochondrial oligoribonuclease, RNA exonuclease 2 (REXO2), has been shown to degrade small single-stranded RNA and DNA oligomers and has been localised to the intermembrane space as well as the matrix (Bruni et al. 2013). The mammalian endonuclease EndoG has been shown to digest a range of different nucleic acid species, including single- and double-stranded RNA (Cymerman et al. 2008), and has also been localised to the intermembrane space, where it is proposed to function in apoptosis (Li et al. 2001). The presence of so many putative ribonuclease in the intermembrane space raises the possibility that degradation can occur there, but what RNAs are being degraded, or whether these enzymes protect mitochondria from cytosolic RNAs, is not yet clear. Thus, although great advances have been made in our understanding of mitochondrial RNA decay, the complexity and seemingly different routes of mtRNA metabolism in different species have made formulating a unifying mechanism impossible. However, a range of recent advances, such as improved proteomic analysis, vastly reduced costs for next-generation sequencing, improved structural analytical tools and novel screening techniques as well as the incredible creativity of the research community, has brought us closer to identifying all of the components and mechanisms involved in mitochondrial RNA metabolism.

2.9 Clinical Relevance

An additional source has come from the growing list of genes involved in mitochondrial diseases. Whole exome or whole genome sequencing is now routinely used as a diagnostic tool to identify monogenic diseases, and to date, more than 500 genes have been associated with inborn errors of metabolism (IEM) alone, many with mitochondrial involvement. The necessity to functionally validate novel disease-causing variants has resulted in the identification of several genes directly involved in mitochondrial RNA metabolism.

Mitochondrial diseases are defined by a primary defect in the OXPHOS system, but mitochondrial dysfunction is also a hallmark in other diseases, such as heart disease, diabetes, cancer, neurodegeneration and even the natural ageing process (Nunnari and Suomalainen 2012). The mitochondrial network is essential in almost all tissues, and thus its dysfunction can result in a wide range of symptoms or combinations of symptoms. Clinical manifestations often involve the brain, heart and skeletal muscle, with a vast spectrum of different clinical phenotypes, such as proximal muscle weakness, cardiomyopathy, endocrine aberrations, external ophthalmoplegia, mental retardation, dementia, ataxia, seizures, stroke-like episodes, deafness and blindness. These diseases can be caused by mutations in the mitochondrial genome or in nuclear genes important for mitochondrial function. Around 100 pathogenic mutations in mtDNA have been confirmed, while the list of nucleus-encoded variants involved in mitochondrial disease is continuously growing. For instance, in POLγ alone, more than 250 mutations have been identified to date (Copeland 2014). Disease-causative genes have now been identified in all processes of maintaining and expressing the mitochondrial genome, including LRPPRC, RNase P, ELAC2, MTPAP, PNPase, mitochondrial aminoacyl-tRNA synthases, NSUN3 and factors involved in the assembly and function of the mitoribosome (Mootha et al. 2003; Crosby et al. 2010; von Ameln et al. 2012; Haack et al. 2013; Vilardo and Rossmanith 2015; Falk et al. 2016; Metodiev et al. 2016; Van Haute et al. 2016).

2.10 Future Questions and Concluding Remarks

Several key questions in mitochondrial RNA metabolism remain. By now it is clear that in mammalian and some metazoan mitochondria, polyadenylation does not strictly serve as a signal for degradation and the possibility of two polyadenylation states is suggested, where a stable poly(A) tail marks transcript maturation, while transient polyadenylation is involved in the decay pathway. Regulating between these two polyadenylation states could be controlled by specific deadenylases, such as the 2′,5′-phosphodiesterase 12, PDE12, which has been shown to be involved in the maturation of some mt-tRNAs (Rorbach et al. 2011; Pearce et al. 2017) or other, not yet characterised deadenylases. We have shown that the poly(A) tail is required

to maintain mt-mRNA integrity and that the polyadenylation process is needed for the maturation of specific mt-tRNAs but not for translation. It is thus not unlikely that polyadenylation has additional functions, depending on tail length and sub-mitochondrial localisation. There are clear differences in the rate of RNA decay, depending on the type of transcript, and although LRPPRC is able to stabilise mRNA, it does not seem to possess sequence specificity, raising the question of what regulates stability. Further, neither tRNAs nor rRNAs seem to be regulated by the proposed ribonucleases, and despite rRNAs having a higher rate of transcription, their turnover is not understood. Thus, elucidating the exact machinery and solving the mechanism will be important to understand mitochondrial biology.

References

Adán C, Matsushima Y, Hernández-Sierra R et al (2008) Mitochondrial transcription factor B2 is essential for metabolic function in *Drosophila melanogaster* development. J Biol Chem 283:12333–12342. https://doi.org/10.1074/jbc.M801342200

Alam TI, Kanki T, Muta T et al (2003) Human mitochondrial DNA is packaged with TFAM. Nucleic Acids Res 31:1640–1645

Allen JF (2017) The CoRR hypothesis for genes in organelles. J Theor Biol 434:50–57. https://doi.org/10.1016/j.jtbi.2017.04.008

Aloni Y, Attardi G (1971) Symmetrical *in vivo* transcription of mitochondrial DNA in HeLa cells. Proc Natl Acad Sci U S A 68:1757–1761

Amaral A, Ramalho-Santos J, St John JC (2007) The expression of polymerase gamma and mitochondrial transcription factor a and the regulation of mitochondrial DNA content in mature human sperm. Hum Reprod 22:1585–1596. https://doi.org/10.1093/humrep/dem030

Amunts A, Brown A, Toots J et al (2015) The structure of the human mitochondrial ribosome. Science 348:95–98. https://doi.org/10.1126/science.aaa1193

Andreenkov OV, Volkova EI, Demakov SA et al (2016) Targeted mutagenesis of *Drosophila RNaseZ* gene by homologous recombination. Dokl Biochem Biophys 471:399–402. https://doi.org/10.1134/S1607672916060065

Antonicka H, Shoubridge EA (2015) Mitochondrial RNA granules are centers for posttranscriptional RNA processing and ribosome biogenesis. Cell Rep 10:920–932. https://doi.org/10.1016/j.celrep.2015.01.030

Antonicka H, Sasarman F, Nishimura T et al (2013) The mitochondrial RNA-binding protein GRSF1 localizes to RNA granules and is required for posttranscriptional mitochondrial gene expression. Cell Metab 17:386–398. https://doi.org/10.1016/j.cmet.2013.02.006

Aphasizhev R, Aphasizheva I (2008) Terminal RNA uridylyltransferases of trypanosomes. Biochim Biophys Acta 1779:270–280. https://doi.org/10.1016/j.bbagrm.2007.12.007

Arroyo JD, Jourdain AA, Calvo SE et al (2016) A genome-wide CRISPR death screen identifies genes essential for oxidative phosphorylation. Cell Metab 24:875–885. https://doi.org/10.1016/j.cmet.2016.08.017

Asin-Cayuela J, Schwend T, Farge G, Gustafsson CM (2005) The human mitochondrial transcription termination factor (mTERF) is fully active *in vitro* in the non-phosphorylated form. J Biol Chem 280:25499–25505. https://doi.org/10.1074/jbc.M501145200

Augustin MA, Reichert AS, Betat H et al (2003) Crystal structure of the human CCA-adding enzyme: insights into template-independent polymerization. J Mol Biol 328:985–994

Baggio F, Bratic A, Mourier A et al (2014) *Drosophila melanogaster* LRPPRC2 is involved in coordination of mitochondrial translation. Nucleic Acids Res 42:13920–13938. https://doi.org/10.1093/nar/gku1132

Bai Y, Srivastava SK, Chang JH et al (2011) Structural basis for dimerization and activity of human PAPD1, a noncanonical poly(a) polymerase. Mol Cell 41:311–320. https://doi.org/10.1016/j.molcel.2011.01.013

Barkan A, Small ID (2014) Pentatricopeptide repeat proteins in plants. Annu Rev Plant Biol 65:415–442. https://doi.org/10.1146/annurev-arplant-050213-040159

Bar-Yaacov D, Frumkin I, Yashiro Y et al (2016) Mitochondrial 16S rRNA is methylated by tRNA methyltransferase TRMT61B in all vertebrates. PLoS Biol 14:e1002557

Battey J, Clayton DA (1978) The transcription map of mouse mitochondrial DNA. Cell 14:143–156

Bernstein JA, Khodursky AB, Lin P-H et al (2002) Global analysis of mRNA decay and abundance in *Escherichia coli* at single-gene resolution using two-color fluorescent DNA microarrays. Proc Natl Acad Sci U S A 99:9697–9702. https://doi.org/10.1073/pnas.112318199

Bilbille Y, Gustilo EM, Harris KA et al (2011) The human mitochondrial $tRNA^{Met}$: structure/function relationship of a unique modification in the decoding of unconventional codons. J Mol Biol 406:257–274. https://doi.org/10.1016/j.jmb.2010.11.042

Boehm E, Zaganelli S, Maundrell K et al (2017) FASTKD1 and FASTKD4 have opposite effects on expression of specific mitochondrial RNAs, depending upon their endonuclease-like RAP domain. Nucleic Acids Res 45:6135–6146. https://doi.org/10.1093/nar/gkx164

Bogenhagen DF, Applegate EF, Yoza BK (1984) Identification of a promoter for transcription of the heavy strand of human mtDNA: in vitro transcription and deletion mutagenesis. Cell 36:1105–1113

Bogenhagen DF, Martin DW, Koller A (2014) Initial steps in RNA processing and ribosome assembly occur at mitochondrial DNA nucleoids. Cell Metab 19:618–629. https://doi.org/10.1016/j.cmet.2014.03.013

Borowski LS, Szczesny RJ (2014) Measurement of mitochondrial RNA stability by metabolic labeling of transcripts with 4-thiouridine. Methods Mol Biol 1125:277–286. https://doi.org/10.1007/978-1-62703-971-0_22

Borowski LS, Dziembowski A, Hejnowicz MS et al (2013) Human mitochondrial RNA decay mediated by PNPase-hSuv3 complex takes place in distinct foci. Nucleic Acids Res 41:1223–1240. https://doi.org/10.1093/nar/gks1130

Bratic A, Larsson N-G (2013) The role of mitochondria in aging. J Clin Invest 123:951–957. https://doi.org/10.1172/JCI64125

Bratic A, Wredenberg A, Grönke S et al (2011) The bicoid stability factor controls polyadenylation and expression of specific mitochondrial mRNAs in *Drosophila melanogaster*. PLoS Genet 7:e1002324. https://doi.org/10.1371/journal.pgen.1002324

Bratic A, Kauppila TES, Macao B et al (2015) Complementation between polymerase- and exonuclease-deficient mitochondrial DNA polymerase mutants in genomically engineered flies. Nat Commun 6:8808. https://doi.org/10.1038/ncomms9808

Bratic A, Clemente P, Calvo-Garrido J et al (2016) Mitochondrial polyadenylation is a one-step process required for mRNA integrity and tRNA maturation. PLoS Genet 12:e1006028. https://doi.org/10.1371/journal.pgen.1006028

Brown TA, Tkachuk AN, Shtengel G et al (2011) Superresolution fluorescence imaging of mitochondrial nucleoids reveals their spatial range, limits, and membrane interaction. Mol Cell Biol 31:4994–5010. https://doi.org/10.1128/MCB.05694-11

Bruni F, Manzari C, Filice M et al (2012) D-MTERF5 is a novel factor modulating transcription in *Drosophila* mitochondria. Mitochondrion 12:492–499. https://doi.org/10.1016/j.mito.2012.06.010

Bruni F, Gramegna P, Oliveira JMA et al (2013) REXO2 is an oligoribonuclease active in human mitochondria. PLoS One 8:e64670. https://doi.org/10.1371/journal.pone.0064670

Brzezniak LK, Bijata M, Szczesny RJ, Stepien PP (2011) Involvement of human ELAC2 gene product in 3′ end processing of mitochondrial tRNAs. RNA Biol 8:616–626. https://doi.org/10.4161/rna.8.4.15393

Butow RA, Zhu H, Perlman P, Conrad-Webb H (1989) The role of a conserved dodecamer sequence in yeast mitochondrial gene expression. Genome 31:757–760

Calvo SE, Clauser KR, Mootha VK (2015) MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. Nucleic Acids Res 44:D1251–D1257. https://doi.org/10.1093/nar/gkv1003

Chandel NS (2015) Evolution of mitochondria as signaling organelles. Cell Metab 22:204–206. https://doi.org/10.1016/j.cmet.2015.05.013

Chang DD, Clayton DA (1984) Precise identification of individual promoters for transcription of each strand of human mitochondrial DNA. Cell 36:635–643

Chateigner-Boutin A-L, Small ID (2011) Organellar RNA editing. Wiley Interdiscip Rev RNA 2:493–506. https://doi.org/10.1002/wrna.72

Chen H-W, Rainey RN, Balatoni CE et al (2006) Mammalian polynucleotide phosphorylase is an intermembrane space RNase that maintains mitochondrial homeostasis. Mol Cell Biol 26:8475–8487. https://doi.org/10.1128/MCB.01002-06

Chen P-L, Chen C-F, Chen Y et al (2013) Mitochondrial genome instability resulting from SUV3 haploinsufficiency leads to tumorigenesis and shortened lifespan. Oncogene 32:1193–1201. https://doi.org/10.1038/onc.2012.120

Chujo T, Ohira T, Sakaguchi Y et al (2012) LRPPRC/SLIRP suppresses PNPase-mediated mRNA decay and promotes polyadenylation in human mitochondria. Nucleic Acids Res 40:8033–8047. https://doi.org/10.1093/nar/gks506

Clemente P, Pajak A, Laine I et al (2015) SUV3 helicase is required for correct processing of mitochondrial transcripts. Nucleic Acids Res 43:7398–7413. https://doi.org/10.1093/nar/gkv692

Conrad-Webb H, Perlman PS, Zhu H, Butow RA (1990) The nuclear SUV3-1 mutation affects a variety of post-transcriptional processes in yeast mitochondria. Nucleic Acids Res 18:1369–1376

Copeland WC (2014) Defects of mitochondrial DNA replication. J Child Neurol 29:1216–1224. https://doi.org/10.1177/0883073814537380

Crosby AH, Patel H, Chioza BA et al (2010) Defective mitochondrial mRNA maturation is associated with spastic ataxia. Am J Hum Genet 87:655–660. https://doi.org/10.1016/j.ajhg.2010.09.013

Cymerman IA, Chung I, Beckmann BM et al (2008) EXOG, a novel paralog of endonuclease G in higher eukaryotes. Nucleic Acids Res 36:1369–1379. https://doi.org/10.1093/nar/gkm1169

Dairaghi DJ, Shadel GS, Clayton DA (1995) Addition of a 29 residue carboxyl-terminal tail converts a simple HMG box-containing protein into a transcriptional activator. J Mol Biol 249:11–28. https://doi.org/10.1006/jmbi.1995.9889

DeLuca SZ, O'Farrell PH (2012) Barriers to male transmission of mitochondrial DNA in sperm development. Dev Cell 22:660–668. https://doi.org/10.1016/j.devcel.2011.12.021

Dmochowska A, Kalita K, Krawczyk M et al (1999) A human putative Suv3-like RNA helicase is conserved between Rhodobacter and all eukaryotes. Acta Biochim Pol 46:155–162

Dubrovsky EB, Dubrovskaya VA, Levinger L et al (2004) *Drosophila* RNase Z processes mitochondrial and nuclear pre-tRNA 3′ ends *in vivo*. Nucleic Acids Res 32:255–262. https://doi.org/10.1093/nar/gkh182

Dziembowski A, Piwowarski J, Hoser R et al (2003) The yeast mitochondrial degradosome. Its composition, interplay between RNA helicase and RNase activities and the role in mitochondrial RNA metabolism. J Biol Chem 278:1603–1611. https://doi.org/10.1074/jbc.M208287200

Ekstrand M, Falkenberg M, Rantanen A et al (2004) Mitochondrial transcription factor a regulates mtDNA copy number in mammals. Hum Mol Genet 13:935–944. https://doi.org/10.1093/hmg/ddh109

el Meziane A, Callen JC, Mounolou JC (1989) Mitochondrial gene expression during *Xenopus laevis* development: a molecular study. EMBO J 8:1649–1655

Falk MJ, Gai X, Shigematsu M et al (2016) A novel HSD17B10 mutation impairing the activities of the mitochondrial RNase P complex causes X-linked intractable epilepsy and neurodevelopmental regression. RNA Biol 13:477–485. https://doi.org/10.1080/15476286.2016.1159381

Falkenberg M, Gaspari M, Rantanen A et al (2002) Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA. Nat Genet 31:289–294. https://doi.org/10.1038/ng909

Falkenberg M, Larsson N-G, Gustafsson CM (2007) DNA replication and transcription in mammalian mitochondria. Annu Rev Biochem 76:679–699. https://doi.org/10.1146/annurev.biochem.76.060305.152028

Fernández-Vizarra E, Berardinelli A, Valente L et al (2007) Nonsense mutation in pseudouridylate synthase 1 (PUS1) in two brothers affected by myopathy, lactic acidosis and sideroblastic anaemia (MLASA). J Med Genet 44:173–180. https://doi.org/10.1136/jmg.2006.045252

Freyer C, Park CB, Ekstrand M et al (2010) Maintenance of respiratory chain function in mouse hearts with severely impaired mtDNA transcription. Nucleic Acids Res 38:6577–6588. https://doi.org/10.1093/nar/gkq527

Garesse R, Kaguni LS (2005) A *Drosophila* model of mitochondrial DNA replication: proteins, genes and regulation. IUBMB Life 57:555–561. https://doi.org/10.1080/15216540500215572

Gohil VM, Nilsson R, Belcher-Timme CA et al (2010) Mitochondrial and nuclear genomic responses to loss of LRPPRC expression. J Biol Chem 285:13742–13747. https://doi.org/10.1074/jbc.M109.098400

Groot GS, Flavell RA, Van Ommen GJ, Grivell LA (1974) Yeast mitochondrial RNA does not contain poly(a). Nature 252:167–169

Gustafsson CM, Falkenberg M, Larsson N-G (2016) Maintenance and expression of mammalian mitochondrial DNA. Annu Rev Biochem 85:133–160. https://doi.org/10.1146/annurev-biochem-060815-014402

Haack TB, Kopajtich R, Freisinger P et al (2013) ELAC2 mutations cause a mitochondrial RNA processing defect associated with hypertrophic cardiomyopathy. Am J Hum Genet 93 (2):211–223. https://doi.org/10.1016/j.ajhg.2013.06.006

Haag S, Sloan KE, Ranjan N et al (2016) NSUN3 and ABH1 modify the wobble position of mt-tRNA^{Met} to expand codon recognition in mitochondrial translation. EMBO J 35:2104–2119. https://doi.org/10.15252/embj.201694885

Harmel J, Ruzzenente B, Terzioglu M et al (2013) The leucine-rich pentatricopeptide repeat-containing protein (LRPPRC) does not activate transcription in mammalian mitochondria. J Biol Chem 288:15510–15519. https://doi.org/10.1074/jbc.M113.471649

Helm M, Brulé H, Degoul F et al (1998) The presence of modified nucleotides is required for cloverleaf folding of a human mitochondrial tRNA. Nucleic Acids Res 26:1636–1643

Helm M, Giegé R, Florentz C (1999) A Watson-Crick base-pair-disrupting methyl group (m^1A9) is sufficient for cloverleaf folding of human mitochondrial tRNA^{Lys}. Biochemistry 38:13338–13346

Hillen HS, Parshin AV, Agaronyan K et al (2017) Mechanism of transcription anti-termination in human mitochondria. Cell 171:1082–1093.e13. https://doi.org/10.1016/j.cell.2017.09.035

Hirayama T, Matsuura T, Ushiyama S et al (2013) A poly(a)-specific ribonuclease directly regulates the poly(a) status of mitochondrial mRNA in *Arabidopsis*. Nat Commun 4:2247. https://doi.org/10.1038/ncomms3247

Hoffmann B, Nickel J, Speer F, Schafer B (2008) The 3′ ends of mature transcripts are generated by a processosome complex in fission yeast mitochondria. J Mol Biol 377:1024–1037. https://doi.org/10.1016/j.jmb.2008.01.038

Houseley J, Tollervey D (2009) The many pathways of RNA degradation. Cell 136:763–776. https://doi.org/10.1016/j.cell.2009.01.019

Hui MP, Foley PL, Belasco JG (2014) Messenger RNA degradation in bacterial cells. Annu Rev Genet 48:537–559. https://doi.org/10.1146/annurev-genet-120213-092340

Hyvärinen AK, Pohjoismäki JLO, Reyes A et al (2007) The mitochondrial transcription termination factor mTERF modulates replication pausing in human mitochondrial DNA. Nucleic Acids Res 35:6458–6474. https://doi.org/10.1093/nar/gkm676

Iborra FJ, Kimura H, Cook PR (2004) The functional organization of mitochondrial genomes in human cells. BMC Biol 2:9. https://doi.org/10.1186/1741-7007-2-9

Iwata S, Lee JW, Okada K et al (1998) Complete structure of the 11-subunit bovine mitochondrial cytochrome bc_1 complex. Science 281:64–71

Jacob ST, Schindler DG (1972) Polyriboadenylate polymerase solubilized from rat liver mitochondria. Biochem Biophys Res Commun 48:126–134

Jemt E, Persson Ö, Shi Y et al (2015) Regulation of DNA replication at the end of the mitochondrial D-loop involves the helicase TWINKLE and a conserved sequence element. Nucleic Acids Res 43:9262–9275. https://doi.org/10.1093/nar/gkv804

Jourdain AA, Koppen M, Wydro M et al (2013) GRSF1 regulates RNA processing in mitochondrial RNA granules. Cell Metab 17:399–410. https://doi.org/10.1016/j.cmet.2013.02.005

Jourdain AA, Koppen M, Rodley CD et al (2015) A mitochondria-specific isoform of FASTK is present in mitochondrial RNA granules and regulates gene expression and function. Cell Rep 10:1110–1121. https://doi.org/10.1016/j.celrep.2015.01.063

Jourdain AA, Boehm E, Maundrell K, Martinou J-C (2016) Mitochondrial RNA granules: compartmentalizing mitochondrial gene expression. J Cell Biol 212:611–614. https://doi.org/10.1083/jcb.201507125

Jourdain AA, Popow J, la Fuente De MA, et al (2017) The FASTK family of proteins: emerging regulators of mitochondrial RNA biology. Nucleic Acids Res45(19):10941-10947. doi: https://doi.org/10.1093/nar/gkx772

Kaguni LS (2004) DNA polymerase γ, the mitochondrial replicase. Annu Rev Biochem 73:293–320. https://doi.org/10.1146/annurev.biochem.72.121801.161455

Kalckar H (1939) The nature of phosphoric esters formed in kidney extracts. Biochem J 33:631–641

Karasik A, Shanmuganathan A, Howard MJ et al (2016) Nuclear protein-only ribonuclease P2 structure and biochemical characterization provide insight into the conserved properties of tRNA 5′ end processing enzymes. J Mol Biol 428:26–40. https://doi.org/10.1016/j.jmb.2015.11.025

Kauppila TES, Kauppila JHK, Larsson N-G (2017) Mammalian mitochondria and aging: An update. Cell Metab 25:57–71

Kolesnikov AA, Gerasimov ES (2012) Diversity of mitochondrial genome organization. Biochemistry (Mosc) 77:1424–1435. https://doi.org/10.1134/S0006297912130020

Kruse B, Narasimhan N, Attardi G (1989) Termination of transcription in human mitochondria: identification and purification of a DNA binding protein factor that promotes termination. Cell 58:391–397

Kühl I, Kukat C, Ruzzenente B et al (2014) POLRMT does not transcribe nuclear genes. Nature 514:E7–E11. https://doi.org/10.1038/nature13690

Kukat C, Larsson N-G (2013) mtDNA makes a U-turn for the mitochondrial nucleoid. Trends Cell Biol 23:457–463. https://doi.org/10.1016/j.tcb.2013.04.009

Kukat C, Wurm CA, Spåhr H et al (2011) Super-resolution microscopy reveals that mammalian mitochondrial nucleoids have a uniform size and frequently contain a single copy of mtDNA. Proc Natl Acad Sci U S A 108:13534–13539. https://doi.org/10.1073/pnas.1109263108

Kukat C, Davies KM, Wurm CA et al (2015) Cross-strand binding of TFAM to a single mtDNA molecule forms the mitochondrial nucleoid. Proc Natl Acad Sci U S A 112:11288–11293. https://doi.org/10.1073/pnas.1512131112

Ladoukakis ED, Zouros E (2017) Evolution and inheritance of animal mitochondrial DNA: rules and exceptions. J Biol Res (Thessalon) 24:2. https://doi.org/10.1186/s40709-017-0060-4

Lagouge M, Mourier A, Lee HJ et al (2015) SLIRP regulates the rate of mitochondrial protein synthesis and protects LRPPRC from degradation. PLoS Genet 11:e1005423. https://doi.org/10.1371/journal.pgen.1005423

Lakshmipathy U, Campbell C (1999) The human DNA ligase III gene encodes nuclear and mitochondrial proteins. Mol Cell Biol 19:3869–3876

Lane N (2017) Serial endosymbiosis or singular event at the origin of eukaryotes? J Theor Biol 434:58–67. https://doi.org/10.1016/j.jtbi.2017.04.031

Lane N, Martin W (2010) The energetics of genome complexity. Nature 467:929–934. https://doi.org/10.1038/nature09486

Lapkouski M, Hallberg BM (2015) Structure of mitochondrial poly(a) RNA polymerase reveals the structural basis for dimerization, ATP selectivity and the SPAX4 disease phenotype. Nucleic Acids Res 43:9065–9075. https://doi.org/10.1093/nar/gkv861

Larsson N-G (2010) Somatic mitochondrial DNA mutations in mammalian aging. Annu Rev Biochem 79:683–706. https://doi.org/10.1146/annurev-biochem-060408-093701

Larsson N-G, Oldfors A, Garman JD et al (1997) Down-regulation of mitochondrial transcription factor a during spermatogenesis in humans. Hum Mol Genet 6:185–191

Larsson N-G, Wang J, Wilhelmsson H et al (1998) Mitochondrial transcription factor a is necessary for mtDNA maintenance and embryogenesis in mice. Nat Genet 18:231–236. https://doi.org/10.1038/ng0398-231

Lee K-W, Bogenhagen DF (2014) Assignment of 2′-O-methyltransferases to modification sites on the mammalian mitochondrial large subunit 16 S ribosomal RNA (rRNA). J Biol Chem 289:24936–24942. https://doi.org/10.1074/jbc.C114.581868

Lee K-W, Okot-Kotber C, LaComb JF, Bogenhagen DF (2013) Mitochondrial ribosomal RNA (rRNA) methyltransferase family members are positioned to modify nascent rRNA in foci near the mitochondrial DNA nucleoid. J Biol Chem 288:31386–31399. https://doi.org/10.1074/jbc.M113.515692

Levin L, Mishmar D (2015) A genetic view of the mitochondrial role in ageing: killing us softly. Adv Exp Med Biol 847:89–106. https://doi.org/10.1007/978-1-4939-2404-2_4

Levy S, Allerston CK, Liveanu V et al (2016) Identification of LACTB2, a metallo-β-lactamase protein, as a human mitochondrial endoribonuclease. Nucleic Acids Res 44:1813–1832. https://doi.org/10.1093/nar/gkw050

Lewis DL, Farr CL, Kaguni LS (1995) *Drosophila melanogaster* Mitochondrial DNA: completion of the nucleotide sequence and evolutionary comparisons. Insect Mol Biol 4:263–278. https://doi.org/10.1111/j.1365-2583.1995.tb00032.x

Li LY, Luo X, Wang X (2001) Endonuclease G is an apoptotic DNase when released from mitochondria. Nature 412:95–99. https://doi.org/10.1038/35083620

Lightowlers RN, Chrzanowska-Lightowlers ZMA (2013) Human pentatricopeptide proteins: only a few and what do they do? RNA Biol 10:1433–1438. https://doi.org/10.4161/rna.24770

Lin CL, Wang Y-T, Yang W-Z et al (2012) Crystal structure of human polynucleotide phosphorylase: insights into its domain function in RNA binding and degradation. Nucleic Acids Res 40:4146–4157. https://doi.org/10.1093/nar/gkr1281

Litonin D, Litonin D, Sologub M et al (2010) Human mitochondrial transcription revisited: only TFAM and TFB2M are required for transcription of the mitochondrial genes *in vitro*. J Biol Chem 285:18129–18133. https://doi.org/10.1074/jbc.C110.128918

Liu P, Huang J, Zheng Q et al (2017) Mammalian mitochondrial RNAs are degraded in the mitochondrial intermembrane space by RNASET2. Protein Cell 8:735–749. https://doi.org/10.1007/s13238-017-0448-9

Ma H, Xu H, O'Farrell PH (2014) Transmission of mitochondrial mutations and action of purifying selection in Drosophila melanogaster. Nat Genet 46:393–397. https://doi.org/10.1038/ng.2919

Macao B, Uhler JP, Siibak T et al (2015) The exonuclease activity of DNA polymerase γ is required for ligation during mitochondrial DNA replication. Nat Commun 6:7303. https://doi.org/10.1038/ncomms8303

Malarkey CS, Bestwick M, Kuhlwilm JE et al (2012) Transcriptional activation by mitochondrial transcription factor a involves preferential distortion of promoter DNA. Nucleic Acids Res 40:614–624. https://doi.org/10.1093/nar/gkr787

Matsushima Y, Matsumura K, Ishii S et al (2003) Functional domains of chicken mitochondrial transcription factor a for the maintenance of mitochondrial DNA copy number in lymphoma cell line DT40. J Biol Chem 278:31149–31158. https://doi.org/10.1074/jbc.M303842200

Matsushima Y, Hirofuji Y, Aihara M et al (2017) *Drosophila* protease ClpXP specifically degrades DmLRPPRC1 controlling mitochondrial mRNA and translation. Sci Rep 7:8315. https://doi.org/10.1038/s41598-017-08088-6

Mercer TR, Neph S, Dinger ME et al (2011) The human mitochondrial transcriptome. Cell 146:645–658. https://doi.org/10.1016/j.cell.2011.06.051

Metodiev MD, Lesko N, Park CB et al (2009) Methylation of 12S rRNA is necessary for in vivo stability of the small subunit of the mammalian mitochondrial ribosome. Cell Metab 9:386–397. https://doi.org/10.1016/j.cmet.2009.03.001

Metodiev MD, Spåhr H, Polosa PL et al (2014) NSUN4 dual function mitochondrial protein required for both methylation of 12S rRNA and coordination of mitoribosomal assembly. PLoS Genet 10:e1004110. https://doi.org/10.1371/journal.pgen.1004110

Metodiev MD, Thompson K, Alston CL et al (2016) Recessive mutations in TRMT10C cause defects in mitochondrial RNA processing and multiple respiratory chain deficiencies. Am J Hum Genet 98:993–1000. https://doi.org/10.1016/j.ajhg.2016.03.010

Mili S, Piñol-Roma S (2003) LRP130, a pentatricopeptide motif protein with a noncanonical RNA-binding domain, is bound in vivo to mitochondrial and nuclear RNAs. Mol Cell Biol 23:4972–4982. https://doi.org/10.1128/MCB.23.14.4972-4982.2003

Minczuk M, He J, He J et al (2011) TEFM (c17orf42) is necessary for transcription of human mtDNA. Nucleic Acids Res 39:4284–4299. https://doi.org/10.1093/nar/gkq1224

Mishra P, Chan DC (2014) Mitochondrial dynamics and inheritance during cell division, development and disease. Nat Rev Mol Cell Biol 15:634–646. https://doi.org/10.1038/nrm3877

Mitchell P (1961) Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. Nature 191:144–148

Montoya J, Christianson TW, Levens D et al (1982) Identification of initiation sites for heavy-strand and light-strand transcription in human mitochondrial DNA. Proc Natl Acad Sci U S A 79:7195–7199

Mootha VK, Lepage P, Miller K et al (2003) Identification of a gene causing human cytochrome *c* oxidase deficiency by integrative genomics. Proc Natl Acad Sci U S A 100:605–610. https://doi.org/10.1073/pnas.242716699

Mourier A, Ruzzenente B, Brandt T et al (2014) Loss of LRPPRC causes ATP synthase deficiency. Hum Mol Genet 23:2580–2592. https://doi.org/10.1093/hmg/ddt652

Nagaike T, Suzuki T, Tomari Y et al (2001) Identification and characterization of mammalian mitochondrial tRNA nucleotidyltransferases. J Biol Chem 276:40041–40049. https://doi.org/10.1074/jbc.M106202200

Nagaike T, Suzuki T, Katoh T, Ueda T (2005) Human mitochondrial mRNAs are stabilized with polyadenylation regulated by mitochondria-specific poly(a) polymerase and polynucleotide phosphorylase. J Biol Chem 280:19721–19727. https://doi.org/10.1074/jbc.M500804200

Nagao A, Hino Shigi N, Suzuki T (2008) Measuring mRNA decay in human mitochondria. Meth Enzymol 447:489–499. https://doi.org/10.1016/S0076-6879(08)02223-4

Nouws J, Goswami AV, Bestwick M et al (2016) Mitochondrial ribosomal protein L12 is required for POLRMT stability and exists as two forms generated by alternative proteolysis during import. J Biol Chem 291:989–997. https://doi.org/10.1074/jbc.M115.689299

Nunnari J, Suomalainen A (2012) Mitochondria: in sickness and in health. Cell 148:1145–1159. https://doi.org/10.1016/j.cell.2012.02.035

Ojala D, Merkel C, Gelfand R, Attardi G (1980) The tRNA genes punctuate the reading of genetic information in human mitochondrial DNA. Cell 22:393–403. https://doi.org/10.1016/0092-8674(80)90350-5

Ojala D, Montoya J, Attardi G (1981) tRNA punctuation model of RNA processing in human mitochondria. Nature 290:470–474. https://doi.org/10.1038/290470a0

Parisi MA, Clayton DA (1991) Similarity of human mitochondrial transcription factor 1 to high mobility group proteins. Science 252:965–969

Parisi MA, Xu B, Clayton DA (1993) A human mitochondrial transcriptional activator can functionally replace a yeast mitochondrial HMG-box protein both *in vivo* and *in vitro*. Mol Cell Biol 13:1951–1961

Patton JR, Bykhovskaya Y, Mengesha E et al (2005) Mitochondrial myopathy and sideroblastic anemia (MLASA): missense mutation in the pseudouridine synthase 1 (PUS1) gene is associated with the loss of pseudouridylation. J Biol Chem 280:19823–19828. https://doi.org/10.1074/jbc.M500216200

Pearce SF, Rebelo-Guiomar P, D'Souza AR et al (2017) Regulation of mammalian mitochondrial gene expression: recent advances. Trends Biochem Sci 42:625–639. https://doi.org/10.1016/j.tibs.2017.02.003

Pham XH, Farge G, Shi Y et al (2006) Conserved sequence box II directs transcription termination and primer formation in mitochondria. J Biol Chem 281:24647–24652. https://doi.org/10.1074/jbc.M602429200

Piechota J, Tomecki R, Gewartowski K et al (2006) Differential stability of mitochondrial mRNA in HeLa cells. Acta Biochim Pol 53:157–168

Popow J, Alleaume A-M, Curk T et al (2015) FASTKD2 is an RNA-binding protein required for mitochondrial RNA processing and translation. RNA 21(11):1873–1884. https://doi.org/10.1261/rna.052365.115

Posse V, Gustafsson CM (2017) Human mitochondrial transcription factor B2 is required for promoter melting during initiation of transcription. J Biol Chem 292:2637–2645. https://doi.org/10.1074/jbc.M116.751008

Posse V, Shahzad S, Falkenberg M et al (2015) TEFM is a potent stimulator of mitochondrial transcription elongation *in vitro*. Nucleic Acids Res 43:2615–2624. https://doi.org/10.1093/nar/gkv105

Puebla-Osorio N, Lacey DB, Alt FW, Zhu C (2006) Early embryonic lethality due to targeted inactivation of DNA ligase III. Mol Cell Biol 26:3935–3941. https://doi.org/10.1128/MCB.26.10.3935-3941.2006

Rackham O, Filipovska A (2012) The role of mammalian PPR domain proteins in the regulation of mitochondrial gene expression. Biochim Biophys Acta 1819:1008–1016. https://doi.org/10.1016/j.bbagrm.2011.10.007

Rackham O, Busch JD, Matic S et al (2016) Hierarchical RNA processing is required for mitochondrial ribosome assembly. Cell Rep 16:1874–1890. https://doi.org/10.1016/j.celrep.2016.07.031

Ramachandran A, Basu U, Sultana S et al (2017) Human mitochondrial transcription factors TFAM and TFB2M work synergistically in promoter melting during transcription initiation. Nucleic Acids Res 45:861–874. https://doi.org/10.1093/nar/gkw1157

Rantanen A, Jansson M, Oldfors A, Larsson N-G (2001) Downregulation of Tfam and mtDNA copy number during mammalian spermatogenesis. Mamm Genome 12:787–792

Reinhard L, Sridhara S, Hallberg BM (2017) The MRPP1/MRPP2 complex is a tRNA-maturation platform in human mitochondria. Nucleic Acids Res 45:12469–12480. https://doi.org/10.1093/nar/gkx902

Ringel R, Sologub M, Morozov YI et al (2011) Structure of human mitochondrial RNA polymerase. Nature 478:269–273. https://doi.org/10.1038/nature10435

Roberti M, Bruni F, Polosa PL et al (2006) The *Drosophila* termination factor DmTTF regulates in vivo mitochondrial transcription. Nucleic Acids Res 34:2109–2116. https://doi.org/10.1093/nar/gkl181

Rorbach J, Nicholls TJJ, Minczuk M (2011) PDE12 removes mitochondrial RNA poly(a) tails and controls translation in human mitochondria. Nucleic Acids Res 39:7750–7763. https://doi.org/10.1093/nar/gkr470

Rorbach J, Boesch P, Gammage PA et al (2014) MRM2 and MRM3 are involved in biogenesis of the large subunit of the mitochondrial ribosome. Mol Biol Cell 25:2542–2555. https://doi.org/10.1091/mbc.E14-01-0014

Rorbach J, Gao F, Powell CA et al (2016) Human mitochondrial ribosomes can switch their structural RNA composition. Proc Natl Acad Sci U S A 113:12198–12201. https://doi.org/10.1073/pnas.1609338113

Rossmanith W (2012) Of P and Z: mitochondrial tRNA processing enzymes. Biochim Biophys Acta 1819:1017–1026. https://doi.org/10.1016/j.bbagrm.2011.11.003

Ruzzenente B, Metodiev MD, Wredenberg A et al (2012) LRPPRC is necessary for polyadenylation and coordination of translation of mitochondrial mRNAs. EMBO J 31:443–456. https://doi.org/10.1038/emboj.2011.392

Safra M, Sas-Chen A, Nir R et al (2017) The m^1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. Nature 12:311. https://doi.org/10.1038/nature24456

Salinas-Giegé T, Cavaiuolo M, Cognat V et al (2017) Polycytidylation of mitochondrial mRNAs in *Chlamydomonas reinhardtii*. Nucleic Acids Res 45:12963–12973. https://doi.org/10.1093/nar/gkx903

Sanchez MIGL, Mercer TR, Davies SMK et al (2011) RNA processing in human mitochondria. Cell Cycle 10:2904–2916. https://doi.org/10.4161/cc.10.17.17060

Sasarman F, Brunel-Guitton C, Antonicka H et al (2010) LRPPRC and SLIRP interact in a ribonucleoprotein complex that regulates posttranscriptional gene expression in mitochondria. Mol Biol Cell 21:1315–1323. https://doi.org/10.1091/mbc.E10-01-0047

Sato M, Sato K (2013) Maternal inheritance of mitochondrial DNA by diverse mechanisms to eliminate paternal mitochondrial DNA. Biochim Biophys Acta 1833:1979–1984. https://doi.org/10.1016/j.bbamcr.2013.03.010

Satoh M, Kuroiwa T (1991) Organization of multiple nucleoids and DNA molecules in mitochondria of a human cell. Exp Cell Res 196:137–140. https://doi.org/10.1016/0014-4827(91)90467-9

Schuster G, Stern D (2009) RNA polyadenylation and decay in mitochondria and chloroplasts. Prog Mol Biol Transl Sci 85:393–422. https://doi.org/10.1016/S0079-6603(08)00810-6

Schwanhäusser B, Busse D, Li N et al (2013) Corrigendum: global quantification of mammalian gene expression control. Nature 495:126–127. https://doi.org/10.1038/nature11848

Seidel-Rogol BL, McCulloch V, Shadel GS (2003) Human mitochondrial transcription factor B1 methylates ribosomal RNA at a conserved stem-loop. Nat Genet 33:23–24. https://doi.org/10.1038/ng1064

Sen A, Karasik A, Shanmuganathan A et al (2016) Loss of the mitochondrial protein-only ribonuclease P complex causes aberrant tRNA processing and lethality in Drosophila. Nucleic Acids Res 44:6409–6422. https://doi.org/10.1093/nar/gkw338

Shafqat N, Marschall H-U, Filling C et al (2003) Expanded substrate screenings of human and Drosophila type 10 17β-hydroxysteroid dehydrogenases (HSDs) reveal multiple specificities in bile acid and steroid hormone metabolism: characterization of multifunctional 3α/7α/7β/17β/20β/21-HSD. Biochem J 376:49–60. doi: https://doi.org/10.1042/BJ20030877

Sheng Z-H, Cai Q (2012) Mitochondrial transport in neurons: impact on synaptic homeostasis and neurodegeneration. Nat Rev Neurosci 13:77–93. https://doi.org/10.1038/nrn3156

Shi Y, Posse V, Zhu X et al (2016) Mitochondrial transcription termination factor 1 directs polar replication fork pausing. Nucleic Acids Res 44:5732–5742. https://doi.org/10.1093/nar/gkw302

Shoubridge EA, Wai T (2007) Mitochondrial DNA and the mammalian oocyte. Curr Top Dev Biol 77:87–111. https://doi.org/10.1016/S0070-2153(06)77004-1

Siira SJ, Spåhr H, Shearwood A-MJ et al (2017) LRPPRC-mediated folding of the mitochondrial transcriptome. Nat Commun 8:1532. https://doi.org/10.1038/s41467-017-01221-z

Sologub M, Litonin D, Anikin M et al (2009) TFB2 is a transient component of the catalytic site of the human mitochondrial RNA polymerase. Cell 139:934–944. https://doi.org/10.1016/j.cell.2009.10.031

Spåhr H, Rozanska A, Li X et al (2016) SLIRP stabilizes LRPPRC via an RRM-PPR protein interface. Nucleic Acids Res 44:6868–6882. https://doi.org/10.1093/nar/gkw575

Stepien PP, Margossian SP, Landsman D, Butow RA (1992) The yeast nuclear gene suv3 affecting mitochondrial post-transcriptional processes encodes a putative ATP-dependent RNA helicase. Proc Natl Acad Sci U S A 89:6813–6817

Sterky FH, Ruzzenente B, Gustafsson CM et al (2010) LRPPRC is a mitochondrial matrix protein that is conserved in metazoans. Biochem Biophys Res Commun 398:759–764. https://doi.org/10.1016/j.bbrc.2010.07.019

Stewart JB, Beckenbach AT (2009) Characterization of mature mitochondrial transcripts in *Drosophila*, and the implications for the tRNA punctuation model in arthropods. Gene 445:49–57. https://doi.org/10.1016/j.gene.2009.06.006

Stewart JB, Chinnery PF (2015) The dynamics of mitochondrial DNA heteroplasmy: implications for human health and disease. Nat Rev Genet 16:530–542. https://doi.org/10.1038/nrg3966

Stewart JB, Larsson N-G (2014) Keeping mtDNA in shape between generations. PLoS Genet 10: e1004670. https://doi.org/10.1371/journal.pgen.1004670

Stewart JB, Freyer C, Elson JL, Larsson N-G (2008) Purifying selection of mtDNA and its implications for understanding evolution and mitochondrial disease. Nat Rev Genet 9:657–662. https://doi.org/10.1038/nrg2396

Stoll B, Zendler D, Binder S (2014) RNA processing factor 7 and polynucleotide phosphorylase are necessary for processing and stability of nad2 mRNA in Arabidopsis mitochondria. RNA Biol 11:968–976. https://doi.org/10.4161/rna.29781

Sun Y, Kurisaki M, Hashiguchi Y, Kumazawa Y (2017) Variation and evolution of polyadenylation profiles in sauropsid mitochondrial mRNAs as deduced from the high-throughput RNA sequencing. BMC Genomics 18:665. https://doi.org/10.1186/s12864-017-4080-0

Sutovsky P, Moreno RD, Ramalho-Santos J et al (1999) Ubiquitin tag for sperm mitochondria. Nature 402:371–372. https://doi.org/10.1038/46466

Suzuki T, Suzuki T (2014) A complete landscape of post-transcriptional modifications in mammalian mitochondrial tRNAs. Nucleic Acids Res 42:7346–7357. https://doi.org/10.1093/nar/gku390

Szczesny RJ, Borowski LS, Brzezniak LK et al (2010) Human mitochondrial RNA turnover caught in flagranti: involvement of hSuv3p helicase in RNA surveillance. Nucleic Acids Res 38:279–298. https://doi.org/10.1093/nar/gkp903

Takemoto C, Spremulli LL, Benkowski LA et al (2009) Unconventional decoding of the AUA codon as methionine by mitochondrial $tRNA^{Met}$ with the anticodon f^5CAU as revealed with a mitochondrial *in vitro* translation system. Nucleic Acids Res 37:1616–1627. https://doi.org/10.1093/nar/gkp001

Temperley RJ, Wydro M, Lightowlers RN, Chrzanowska-Lightowlers ZMA (2010) Human mitochondrial mRNAs—like members of all families, similar but different. Biochim Biophys Acta 1797:1081–1085. https://doi.org/10.1016/j.bbabio.2010.02.036

Terzioglu M, Ruzzenente B, Harmel J et al (2013) MTERF1 binds mtDNA to prevent transcriptional interference at the light-strand promoter but is dispensable for rRNA gene transcription regulation. Cell Metab 17:618–626. https://doi.org/10.1016/j.cmet.2013.03.006

Tomecki R, Dmochowska A, Gewartowski K et al (2004) Identification of a novel human nuclear-encoded mitochondrial poly(a) polymerase. Nucleic Acids Res 32:6001–6014. https://doi.org/10.1093/nar/gkh923

Torregrosa-Muñumer R, Forslund JME, Goffart S et al (2017) PrimPol is required for replication reinitiation after mtDNA damage. Proc Natl Acad Sci U S A 114:11398–11403. https://doi.org/10.1073/pnas.1705367114

Tsukihara T, Aoyama H, Yamashita E et al (1995) Structures of metal sites of oxidized bovine heart cytochrome c oxidase at 2.8 Å. Science 269:1069–1074. https://doi.org/10.1126/science.7652554

Tu Y-T, Barrientos A (2015) The human mitochondrial DEAD-box protein DDX28 resides in RNA granules and functions in mitoribosome assembly. Cell Rep 10:854–864. https://doi.org/10.1016/j.celrep.2015.01.033

Uhler JP, Thörn C, Nicholls TJJ et al (2016) MGME1 processes flaps into ligatable nicks in concert with DNA polymerase γ during mtDNA replication. Nucleic Acids Res 44:5861–5871. https://doi.org/10.1093/nar/gkw468

Van Haute L, Pearce SF, Powell CA et al (2015) Mitochondrial transcript maturation and its disorders. J Inherit Metab Dis 38:655–680. https://doi.org/10.1007/s10545-015-9859-z

Van Haute L, Dietmann S, Kremer L et al (2016) Deficient methylation and formylation of mt-tRNA^{Met} wobble cytosine in a patient carrying mutations in NSUN3. Nat Commun 7:12039. https://doi.org/10.1038/ncomms12039

Vilardo E, Rossmanith W (2015) Molecular insights into HSD10 disease: impact of SDR5C1 mutations on the human mitochondrial RNase P complex. Nucleic Acids Res 43:5112–5119. https://doi.org/10.1093/nar/gkv408

Vilardo E, Nachbagauer C, Buzet A et al (2012) A subcomplex of human mitochondrial RNase P is a bifunctional methyltransferase—extensive moonlighting in mitochondrial tRNA biogenesis. Nucleic Acids Res 40:11583–11593. https://doi.org/10.1093/nar/gks910

von Ameln S, Wang G, Boulouiz R et al (2012) A mutation in PNPT1, encoding mitochondrial-RNA-import protein PNPase, causes hereditary hearing loss. Am J Hum Genet 91:919–927. https://doi.org/10.1016/j.ajhg.2012.09.002

Wai T, Ao A, Zhang X et al (2010) The role of mitochondrial DNA copy number in mammalian fertility. Biol Reprod 83:52–62. https://doi.org/10.1095/biolreprod.109.080887

Wang DD-H, Shu Z, Lieser SA et al (2009) Human mitochondrial SUV3 and polynucleotide phosphorylase form a 330-kDa heteropentamer to cooperatively degrade double-stranded RNA with a 3′-to-5′ directionality. J Biol Chem 284:20812–20821. https://doi.org/10.1074/jbc.M109.009605

Wang G, Chen H-W, Oktay Y et al (2010) PNPASE regulates RNA import into mitochondria. Cell 142:456–467. https://doi.org/10.1016/j.cell.2010.06.035

Wang DD-H, Guo XE, Modrek AS et al (2014) Helicase SUV3, polynucleotide phosphorylase, and mitochondrial polyadenylation polymerase form a transient complex to modulate mitochondrial mRNA polyadenylated tail lengths in response to energetic changes. J Biol Chem 289:16727–16735. https://doi.org/10.1074/jbc.M113.536540

Wanrooij PH, Uhler JP, Simonsson T et al (2010) G-quadruplex structures in RNA stimulate mitochondrial transcription termination and primer formation. Proc Natl Acad Sci U S A 107:16072–16077. https://doi.org/10.1073/pnas.1006026107

Wanrooij PH, Uhler JP, Shi Y et al (2012) A hybrid G-quadruplex structure formed between RNA and DNA explains the extraordinary stability of the mitochondrial R-loop. Nucleic Acids Res 40:10334–10344. https://doi.org/10.1093/nar/gks802

Weill L, Belloc E, Bava F-A, Méndez R (2012) Translational control by changes in poly(a) tail length: recycling mRNAs. Nat Struct Mol Biol 19:577–585. https://doi.org/10.1038/nsmb.2311

Wiedemann N, Pfanner N (2017) Mitochondrial machineries for protein import and assembly. Annu Rev Biochem 86:685–714. https://doi.org/10.1146/annurev-biochem-060815-014352

Wilson WC, Hornig-Do HT, Bruni F et al (2014) A human mitochondrial poly(a) polymerase mutation reveals the complexities of post-transcriptional mitochondrial gene expression. Hum Mol Genet 23:6345–6355. https://doi.org/10.1093/hmg/ddu352

Wolf AR, Mootha VK (2014) Functional genomic analysis of human mitochondrial RNA processing. Cell Rep 7:918–931. https://doi.org/10.1016/j.celrep.2014.03.035

Wolff JN, Nafisinia M, Sutovsky P, Ballard JWO (2013) Paternal transmission of mitochondrial DNA as an integral part of mitochondrial inheritance in metapopulations of *Drosophila simulans*. Heredity (Edinb) 110:57–62. https://doi.org/10.1038/hdy.2012.60

Wydro M, Bobrowicz A, Temperley RJ et al (2010) Targeting of the cytosolic poly(a) binding protein PABPC1 to mitochondria causes mitochondrial translation inhibition. Nucleic Acids Res 38:3732–3742. https://doi.org/10.1093/nar/gkq068

Xie X, Dubrovsky EB (2015) Knockout of *Drosophila* RNase ZL impairs mitochondrial transcript processing, respiration and cell cycle progression. Nucleic Acids Res 43:10364–10375. https://doi.org/10.1093/nar/gkv1149

Xu B, Clayton DA (1995) A persistent RNA-DNA hybrid is formed during transcription at a phylogenetically conserved mitochondrial DNA sequence. Mol Cell Biol 15:580–589

Xu B, Clayton DA (1996) RNA-DNA hybrid formation at the human mitochondrial heavy-strand origin ceases at replication start sites: an implication for RNA-DNA hybrids serving as primers. EMBO J 15:3135–3143

Xu F, Ackerley C, Maj MC et al (2008) Disruption of a mitochondrial RNA-binding protein gene results in decreased cytochrome *b* expression and a marked reduction in ubiquinol-cytochrome c reductase activity in mouse heart mitochondria. Biochem J 416:15–26. https://doi.org/10.1042/BJ20080847

Xu F, Addis JBL, Cameron JM, Robinson BH (2012) LRPPRC mutation suppresses cytochrome oxidase activity by altering mitochondrial RNA transcript stability in a mouse model. Biochem J 441:275–283. https://doi.org/10.1042/BJ20110985

Yakubovskaya E, Mejia E, Byrnes J et al (2010) Helix unwinding and base flipping enable human MTERF1 to terminate mitochondrial transcription. Cell 141:982–993. https://doi.org/10.1016/j.cell.2010.05.018

Yang S-Y, He X-Y, Isaacs C et al (2014) Roles of 17β-hydroxysteroid dehydrogenase type 10 in neurodegenerative disorders. J Steroid Biochem Mol Biol 143:460–472. https://doi.org/10.1016/j.jsbmb.2014.07.001

Young MJ, Copeland WC (2016) Human mitochondrial DNA replication machinery and disease. Curr Opin Genet Dev 38:52–62. https://doi.org/10.1016/j.gde.2016.03.005

Zaganelli S, Rebelo-Guiomar P, Maundrell K et al (2017) The pseudouridine synthase RPUSD4 is an essential component of mitochondrial RNA granules. J Biol Chem 292:4519–4532. https://doi.org/10.1074/jbc.M116.771105

Zhu H, Conrad-Webb H, Liao XS et al (1989) Functional expression of a yeast mitochondrial intron-encoded protein requires RNA processing at a conserved dodecamer sequence at the 3′ end of the gene. Mol Cell Biol 9:1507–1512

Zhu J, Vinothkumar KR, Hirst J (2016) Structure of mammalian respiratory complex I. Nature 536:354–358. https://doi.org/10.1038/nature19095

Zimmer SL, Schein A, Zipor G et al (2009) Polyadenylation in Arabidopsis and *Chlamydomonas* organelles: the input of nucleotidyltransferases, poly(a) polymerases and polynucleotide phosphorylase. Plant J 59:88–99. https://doi.org/10.1111/j.1365-313X.2009.03853.x

Chapter 3
Mitochondrial RNase P Complex in Animals: Mitochondrial tRNA Processing and Links to Disease

Maithili Saoji and Rachel T. Cox

Abstract Loss of mitochondrial function not only causes specific mitochondrial diseases but also contributes to serious conditions such as neurodegeneration and diabetes. Since mitochondrial DNA is transcribed as a polycistronic message comprised of three forms of RNA (rRNA, mRNA, and tRNA), proper 5′- and 3′-end cleavage is essential. In the nucleus, tRNA 5′-end processing is carried out by the first identified ribozyme, RNase P. In contrast, mitochondrial tRNAs are processed by a three-protein complex, mitochondrial RNase P, which does not have an RNA component. An accessory subcomplex made of the m^1A9 methyltransferase MRPP1 and the dehydrogenase MRPP2 binds to the metallonuclease MRPP3 that cleaves the RNA phosphodiester backbone. Each protein has been shown to be essential in model organisms, and loss of each gives rise to human multisystemic diseases with many characteristics of mitochondrial disease. In this review, we discuss what is known about the mitochondrial RNase P complex, the molecular mechanism of 5′-end mitochondrial tRNA processing, and how loss of this activity causes human disease.

3.1 The Polycistronic Nature of Mitochondrial DNA Transcription

3.1.1 The Mitochondrial Genome

Mitochondria have evolved from alphaproteobacteria that resided within the eukaryotic cell, maintaining an endosymbiotic relationship (Gray 2012; Margulis 1970). Owing to their bacterial origin, mitochondria have their own genetic material, but during evolution, most of the bacterial genes were either lost or transferred to the nucleus, with the actual mitochondrial DNA (mtDNA) being much smaller, about

M. Saoji · R. T. Cox (✉)
Department of Biochemistry and Molecular Biology, The Collaborative Health Initiative Research Program (CHIRP), Uniformed Services University, Bethesda, MD, USA
e-mail: rachel.cox@usuhs.edu

J. Cruz-Reyes, M. Gray (eds.), *RNA Metabolism in Mitochondria*, Nucleic Acids and Molecular Biology 34, https://doi.org/10.1007/978-3-319-78190-7_3

16.6 kb, as compared to the bacterial genome (>5 Mb) (Gray 2012; Land et al. 2015). Even though the mitochondrial genome is replicated and maintained independently of the nuclear DNA, numerous nucleus-encoded proteins are required for replication, transcription, posttranscriptional RNA processing, and mitochondrial translation. Around 200–300 nucleus-encoded proteins are translocated into the mitochondrion to bring about mitochondrial gene expression (Powell et al. 2015).

3.1.2 The Punctuation Model

Mitochondrial DNA (mtDNA) is highly conserved across metazoans, encoding the same products often organized differently on the two (heavy and light) strands (Fig. 3.1). *Drosophila* mtDNA is similar to human in its size and with respect to the products it encodes, making *Drosophila* a good model to study human mitochondrial function and disease (Lewis et al. 1995; Sen and Cox 2017) (Fig. 3.1b). mtDNA exercises a complete economy of organization, with no introns, only one intergenic sequence, and smaller rRNAs and tRNAs as compared to their nuclear counterparts. mtDNA encodes 13 proteins that are components of Complexes I

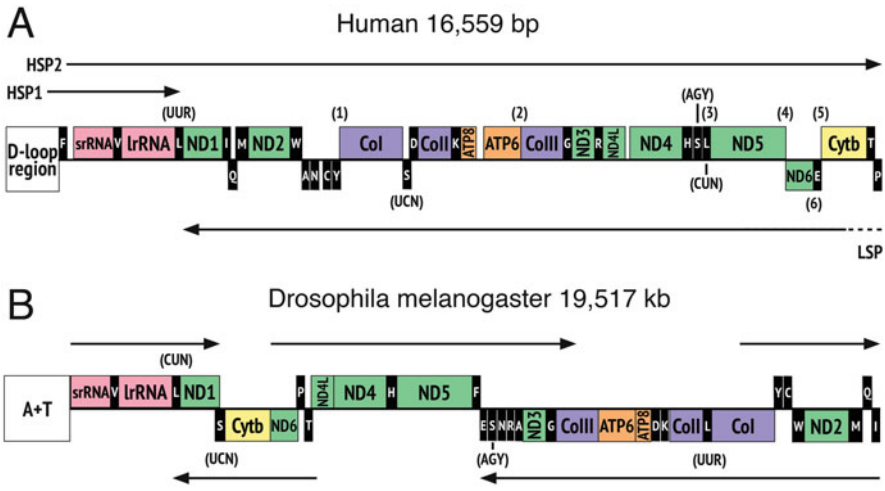

Fig. 3.1 Human and *Drosophila* mtDNA. (**a**) Human mtDNA. Three polycistronic messages are transcribed for human mtDNA (arrows): two promoters initiate on the heavy strand (HSP, top) and one on the light strand (LSP, bottom). With the exception of ND6, all mRNAs and both rRNAs are encoded on the heavy strand. All mt-RNAs are flanked 5′ and 3′ by mt-tRNAs except for the following noncanonical junctions: (1) 5′-end of CoI, (2) ATP8/6 – CoIII, (3) mt-tRNA $^{Leu(CUN)}$—ND5, (4) 3′-end of ND5, (5) 5′-end of Cyt*b*, (6) 3′-end of ND6. (**b**) *Drosophila* mtDNA. The combined five polycistronic messages (arrows) transcribe the same products as human mtDNA. The expanded length is due to the A + T-rich region (not to scale). Single letters (black boxes) represent mt-tRNAs. ND, NADH dehydrogenase (Complex I); Cyt*b*, Cytochrome *b* (Complex III); Co, cytochrome *c* oxidase (Complex IV); and ATP, ATP synthase (Complex V)

(NADH dehydrogenase), III (cytochrome *b*), IV (cytochrome *c* oxidase), and V (ATP synthase) of the OXPHOS pathway. It also encodes 2 ribosomal RNAs (rRNAs) and the complete suite of 22 transfer RNAs (tRNAs) required for the translation of these mitochondria-encoded proteins.

To meet the continuous energy demands of the cells, correct transcription and translation of mitochondria-encoded peptides are obligatory. As in bacteria, mtDNA is transcribed as a polycistronic message. The arrangement and the number of transcripts differ in various species. In humans, mtDNA is transcribed as three polycistronic transcripts, two from the heavy and one from the light strand, whereas in *Drosophila* mtDNA encodes five polycistronic transcripts, three on the heavy and two on the light strand (Fig. 3.1). In humans, the heavy-strand transcripts are considered more informationally rich as they together encode 12 proteins, 2 rRNAs, and 14 tRNAs, while the light-strand transcript encodes 8 tRNAs and only 1 protein (Taanman 1999). The mature RNA products are formed from the polycistronic transcripts following endonucleolytic cleavages at the 5′- and 3′-ends. In most pre-RNA transcripts, the junctions between mitochondrial rRNAs (mt-rRNAs) and mitochondrial mRNAs (mt-mRNAs) are punctuated by tRNAs (Ojala et al. 1981). The secondary cloverleaf structure of mitochondrial tRNAs (mt-tRNAs) is thought to define the cleavage sites for endonucleases to release the individual mature RNA products. Cleavage at the 5′-end of the mt-tRNA molecule is brought about by the three-protein mitochondrial RNase P complex (mtRNase P; Holzmann et al. 2008), whereas a ribonuclease Z (RNase Z)-like protein, ELAC2 in humans and RNase Z^L in *Drosophila*, is responsible for the cleavage at the 3′-end of the mt-tRNAs (Brzezniak et al. 2011; Sanchez et al. 2011).

Most noncanonical mt-mRNAs are flanked by tRNAs [Fig. 3.1a, (1)–(6)]. For 5′-end cleavage of cytochrome *c* oxidase I (CoI), there seems to be a noncoding RNA which adopts a tRNA-like conformation that is recognized and cleaved by mtRNase P (Mercer et al. 2011; Sanchez et al. 2011). The NADH dehydrogenase 5 (ND5)-cytochrome *b* (Cyt *b*) junction is known to be processed by PTCD2, a pentatricopeptide repeat protein (Xu et al. 2008), whereas the processing of the NADH dehydrogenase 6 (ND6)-noncoding (nc) RNA and the ATP6-cytochrome *c* oxidase III (CoIII) junctions is still not clear. Recently, FASTKD5 was shown to be necessary for processing the 5′-end of CoI, ATP8/6-CoIII, and ND5-Cyt*b* junctions, but the exact mechanism is still not well understood (Antonicka and Shoubridge 2015).

3.2 RNase P Function in Mitochondria

3.2.1 Mitochondrial RNase P Cleaves the 5′-End of Mt-tRNA

RNase P is the endonuclease that catalyzes hydrolysis of the phosphodiester bond at the 5′-end of pre-tRNAs to generate tRNAs with a mature 5′-end and a 5′-leader sequence (Robertson et al. 1972). RNase P is found in all domains of life and in

nearly all species. It is a ribozyme with a catalytically active RNA subunit and a variable number of proteins depending on the organism (Klemm et al. 2016). In eukaryotic land plants, algae, and protists, single-peptide protein-only ribonuclease P proteins (PRORPs) are sufficient to bring about 5′-end pre-tRNA processing (Gobert et al. 2010; Klemm et al. 2016; Taschner et al. 2012). In *Arabidopsis*, PRORPs carry out this function in the nucleus, chloroplasts, and mitochondria. However, in human mitochondria, this reaction is carried out by a three-protein mtRNase P complex made up of mitochondrial RNase P protein 1 (MRPP1), mitochondrial RNase P protein 2 (MRPP2), and mitochondrial RNase P protein 3 (MRPP3). In this three-protein complex, MRPP3 is also sometimes referred to as protein-only ribonuclease P (PRORP) (Holzmann et al. 2008).

Unlike its nuclear ribozyme counterpart, mtRNase P lacks a catalytic RNA. Instead, MRPP3, a metallonuclease-like protein, is the catalytic endonuclease that performs the 5′-end cleavage of mitochondrial pre-tRNAs. MRPP1 and 2 have each been recruited from other biological pathways during evolution. MRPP1, or TRMT10C/RG9MTD1, is a methyltransferase required for methylation of the ninth position of certain tRNAs (Jackman et al. 2003), and MRPP2, also known as HDS10 or SDR5C1, is a member of the short-chain dehydrogenase/reductase family which is involved in isoleucine and lipid metabolism (Moeller and Adamski 2009; Shafqat et al. 2003; Yang et al. 2014). MRPP2 acts as a scaffold to form a MRPP1/2 subcomplex and is required for the methyltransferase activity of MRPP1. The catalytic activity of MRPP3 is independent of the methyltransferase activity of the MRPP1/2 subcomplex (Vilardo et al. 2012). The MRPP1/2 subcomplex increases the affinity of MRPP3 for its substrate, likely through a structural rearrangement of mt-tRNA or by altering the active site of MRPP3 (Shafqat et al. 2003). Recently, Reinhard et al. showed that MRPP1/2 subcomplex plays a more central role in coordinating the activities of mt-tRNA processing. They observed that the mt-tRNAs remained bound to the subcomplex after the 5′-end cleavage and that the subcomplex enhances the 3′-end cleavage activity by ELAC2. Furthermore, they observed that MRPP1/2 remained complexed with mt-tRNAs before and after 3′–CCA addition, orchestrating the events leading to the formation of mature mt-tRNA formation (Reinhard et al. 2017).

3.2.2 What Is Known About mtRNase P Mechanism In Vivo

The *Drosophila* homologs of mtRNase P, called Roswell (MRPP1), Scully (MRPP2), and Mulder (MRPP3), have been identified and characterized (Sen et al. 2016). These three proteins share a high degree of sequence similarity with human mtRNase P proteins, containing all the same recognizable domains (Fig. 3.2). They also appear to function in the same way (see below). Sen et al. showed that mutations or knockdown of Mulder (Mldr), Roswell (Rswl), or Scully (Scu) causes lethality in *Drosophila*. Loss or knockdown of each component disrupts mt-RNA processing. Northern blots probed with four mt-tRNAs, each in a different transcript

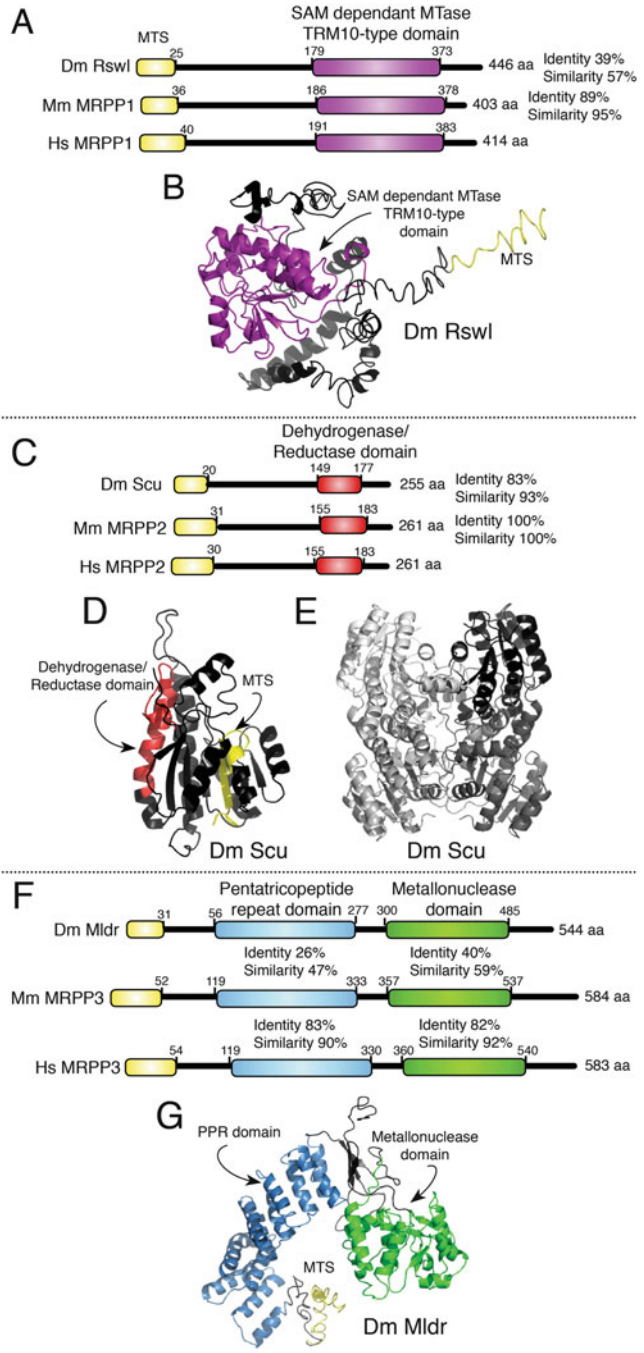

Fig. 3.2 Mitochondrial RNase P complex homologies. Schematics representing the domains of MRPP1, MRPP2, and MRPP3 from *Drosophila*, mouse, and humans are shown in (**a**), (**c**), and (**f**),

environment, showed the accumulation of larger mt-RNA species in mutants. Additionally, overexpressing Mldr and Rswl caused larval lethality with highly reduced levels of ATP as compared to wild type. Lethality was not observed upon overexpressing Scu. Ectopically increasing one of the enzymes of this multiprotein complex could be a dominant negative effect due to sequestration of essential components or mis-regulation of the active enzyme complex formation. Whether the lethality was associated with mis-processing of mt-tRNA remains to be determined.

Recently, MRPP3 was shown to be essential in mouse (see below; Rackham et al. 2016). Using tissues from conditional MRPP3 knockout mice, Northern blots, and q-RT PCR analysis, Rackham et al. showed that MRPP3 loss led to the loss of 5′-end processing of the precursor RNA, resulting in the depletion of mature mt-mRNA over time and to the accumulation of unprocessed, higher molecular weight transcripts. Additionally, they observed an increase in the rate of mtDNA transcription accompanied by a decrease in the synthesis of mitochondria-encoded respiratory complex proteins. The increase in transcription was most likely needed to compensate for the lack of proteins but was unable to rescue the deficient RNA processing. Rackham et al. also demonstrated that previously known canonical RNA junctions depended on MRPP3 for their 5′-end processing. However, the noncanonical junctions, like ND5, which is bordered by $tRNA^{Leu(CUN)}$ at its 5′-end and a noncoding RNA on its 3′-end, are dependent on cleavage by ELAC2 exclusively. Using Northern blot analysis, they demonstrated that the loss of MRPP3 did not have a significant effect on the overall levels of mature ND5 mRNA over time, suggesting that ND5 mRNA maturation is in fact independent of MRPP3 cleavage. Similar independence was also observed for ND6-$tRNA^{Glu}$ and ATP8/6-CoIII junctions, which require ELAC2 and FASTKD5, respectively, for maturation (Sanchez et al. 2011). Interestingly, for some junctions, ELAC2 alone was unable to process the 3′-ends of tRNAs in the absence of 5′-end processing by MRPP3, which was shown by a decrease in the processing at $tRNA^{Val}$-16S rRNA, $tRNA^{Leu(UUR)}$-ND1, and $tRNA^{Met}$-ND2 junctions. This overall pattern of the endonucleolytic cleavages suggests that mt-tRNA maturation happens in a sequential manner, where in most

Fig. 3.2 (continued) respectively. The N-terminal mitochondrial targeting peptide (MTS, yellow) was predicted using the MitoProt server (https://ihg.gsf.de/ihg/mitoprot.html; Claros and Vincens 1996). The domain boundaries are predicted through homology-based alignments performed using Clustal Omega and ExPASy (PROSITE) servers. The percent identity and similarity of individual domains as compared to the human counterparts were calculated using BLAST. The *Drosophila* Roswell (**b**), Scully (**d**, **e**), and Mulder (**g**) protein structural models were created using the I-TASSER server (https://zhanglab.ccmb.med.umich.edu/I-TASSER/, Zhang 2008). The Roswell model is based on the structure of tRNA m^1G9 methyltransferase Trm10 from *Schizosaccharomyces pombe* (PDB ID: 4jwf), whereas the modeled Scully structure is based on human MRPP2/HSD10 (PDB ID: 1u7t). The expected *Drosophila* Scully tetramer assembly is shown in (**e**). The Mulder model was based on *Arabidopsis thaliana* PRORP1, PRORP2, and human MRPP3 structures (PDB IDs: 4G24, 4G23, 5DIZ, 4XGL)

cases 5′-end processing is required for 3′-end cleavage to happen. Furthermore, Rackham et al. also demonstrated that in MRPP3 knockout mice, the expression levels of nucleus-encoded mitochondrial proteins involved in mitochondrial gene expression and mitochondrial ribosome biogenesis were affected. Taken together their results suggested that 5′-end processing by MRPP3 is an indispensable step in mitochondrial RNA maturation, translation, and correct mitoribosome biogenesis.

3.3 RNA Processing Centers in Mitochondria

3.3.1 Mitochondrial RNA Processing Granules

Discrete foci containing newly transcribed mt-RNA were first observed a decade ago by pulse-labeling cells with bromouridine (BrdU). However, the other components of these granules were largely unknown until recently (Fig. 3.3; Iborra et al. 2004). G-rich sequence factor 1 isoform (GRSF1), an RNA-binding protein that binds the G-rich sequence motif AGGGD (where D is either A, U or G), has been previously implicated in posttranscriptional processing of cellular and viral mRNAs (Jablonski and Caputi 2009; Kash et al. 2002; Schaub et al. 2007). GRSF1 was recently shown to accumulate in most BrdU-labeled granules in the mitochondrial matrix (Antonicka et al. 2013; Jourdain et al. 2013). GRSF1 was also shown to play a role in RNA processing even though the exact role of the protein remains unclear. Downregulation of GRSF1 causes a delay in clearing RNA from the mitochondrial RNA granules and affects the steady-state levels of mature RNA transcripts, with an associated decrease in mitochondrial protein levels. It has also been implicated in mitochondrial ribosome biogenesis and stability. Through affinity purification, Jourdain et al. identified several interacting proteins of FLAG-GRSF1 in HEK293T cells. MRPP1 and 2 were the top hits and this interaction was not mediated by RNA. In contrast, while MRPP3 was not identified as a binding partner for GRSF1 through affinity purification, C-terminally tagged MRPP3 was shown to co-localize with GRSF1 in mitochondrial granules using immunocytochemistry (Jourdain et al. 2013). The GRSF1-rich granules were shown to localize newly synthesized precursor RNA transcripts, thus potentially representing sites for mitochondrial RNA processing (Antonicka et al. 2013; Jourdain et al. 2013). In addition, the granules were sensitive to the presence of transcription inhibitors. Jourdain et al. did not observe accumulation of RNase Z in the granules, which could support the sequential cleavage of the 5′- and 3′-ends of mt-tRNAs by mtRNase P and ELAC2, respectively, with the argument that only the first step of mt-tRNA processing happens in the RNA granules.

Nucleoids are the centers in which the replication and transcription of the mitochondrial genome take place. mtRNase P complex proteins and ELAC2 were shown to accumulate in distinct foci adjacent to the nucleoids in mouse 3 T3 cells, and some studies propose that RNA processing granules are extensions of dynamic nucleoids (Borowski et al. 2013). mtRNase P and ELAC2 proteins were shown to be

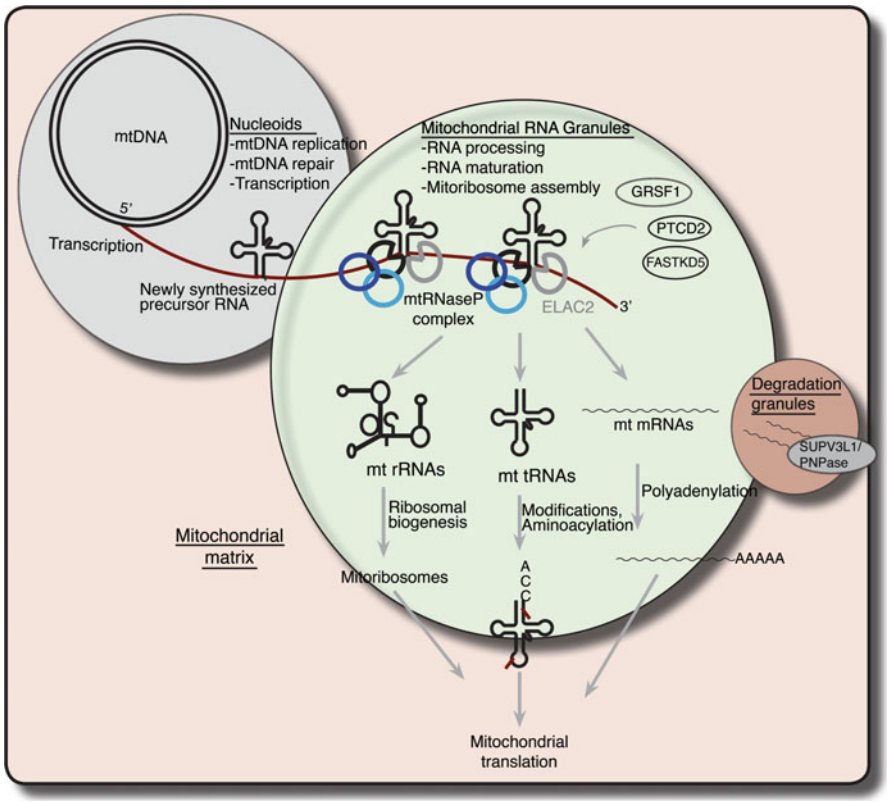

Fig. 3.3 RNA processing granules. mtDNA replication and repair take place in nucleoids. The newly synthesized mt-RNA is processed and matured in mitochondrial RNA granules. The mtRNase P complex performs 5′-end cleavage, and ELAC2 performs 3′-end cleavage of the polycistronic message. Polyadenylation, mt-tRNA modification, and mitoribosome biogenesis also take place in the RNA granules. Mis-processed mt-RNAs are degraded in the associated degradation granules

enriched in the nucleoid granule fractions, even though they are not always co-localized within the nucleoid as judged by immunofluorescence. Thus, it seems likely that the processing of the newly emerged RNA transcript happens either within or very close to the nucleoid structure. Borowski et al. demonstrated a small proportion of the foci co-localize with SUPV3L1/PNPase complex proteins to form a mitochondrial degradosome (Borowski et al. 2013) (Fig. 3.3). As the RNA processing granule research matures, many questions remain regarding the exact order of events. More specifically, does GRSF1 act upstream or downstream of the mtRNase P complex? How would mtRNase P protein inhibition affect granule formation? Will mtRNase P protein downregulation lead to GRSF1 accumulation in mitochondrial RNA processing granules?

3.4 Mitochondrial tRNA Processing and Disease

Since mitochondria play a fundamentally important role in tissues, mutations in mt-tRNAs and defects in mt-tRNA processing lead to diseases that are often severe, sometimes fatal. Currently, treatments are not very effective and there are no cures for these diseases. One unusual trait of mitochondrial diseases is that different mutations in the same mt-tRNA can give rise to very different clinical presentations (Chinnery 2000). There are approximately 275 disease-associated mutations in mt-tRNAs, which often do not affect the anticodon region (Brandon et al. 2005). Mitochondrial diseases can also occur due to defects in the nucleus-encoded enzymes required to carry out the up to 20 mt-tRNA modifications needed to produce mature and fully functional mt-tRNAs (Van Haute et al. 2015).

In metazoans, the full-suite of mt-tRNAs (20) is encoded in mtDNA. However, in contrast to nuclear tRNAs, there are only 2 sets of mitochondrial isoacceptors (different tRNAs species that carry the same amino acids but have different anticodon sequences), whereas there are 513 identified tRNA-encoding genes in the human genome. Thus in mitochondria, there is only 1 mt-tRNA for each of 18 of the amino acids and 2 for serine and leucine. This means that for the majority of mt-tRNAs, if a mt-tRNA becomes nonfunctional due to mutation, the corresponding amino acid will simply not be available for translation. This situation can have differential effects on mt-mRNAs depending on the amino acid requirements for their synthesis.

Due to the polycistronic nature of mtDNA and the mt-tRNA "punctuation model," teasing out the effect of abnormal mt-tRNA levels is challenging. Loss of normal mt-tRNA processing can lead to a decrease in mt-tRNA levels, mitochondrial protein levels, and mt-rRNAs. If available, fully functional mt-tRNAs are too scarce, translation will not occur at normal levels. For example, translation is reduced if enzymes that carry out specific mt-tRNA modifications are defective. However, if mt-tRNAs are not cleaved properly from the polycistronic message, this will have detrimental effects on mt-mRNA processing and mt-rRNA abundance irrespective of mt-tRNA levels. Diseases arising from mt-tRNA processing defects can therefore be classified into two groups: those that arise from mutations in mt-tRNAs and those due to mutations in the nucleus-encoded processing enzymes.

3.4.1 Point Mutations in mt-tRNAs Affecting mt-RNA Processing

Various point mutations in mt-tRNAs have been shown to alter 5′- and 3′-end processing (Table 3.1). The difference thus far appears to be that the mutation must be at the mtRNase P site in order for 5′-end processing to be disrupted, whereas

Table 3.1 mt-tRNA mutations affecting 5′- and 3′-end processing

Mutation[a]	Affected mt-tRNA	Human disease phenotype	Citation(s)
5′-end processing			
Known processing defects			
4263 A → G	tRNAIle	Maternally inherited hypertension	Wang et al. (2011)
5655 A → G	tRNAAla		Jiang et al. (2016)
Hypothesized processing defects			
4401 A → G	$tRNA^{Gln}$ and $tRNA^{Met}$	Maternally inherited hypertension	Li et al. (2009) and Zhu et al. (2009)
5512 A → G	$tRNA^{Trp}$		Guo et al. (2016)
3′-end processing			
Known processing defects			
4269 A → G	$tRNA^{Ile}$	Cardiomyopathy	Levinger et al. (2003)
4295 A → G			
4317 A → G			
4309 G → A[b]		Ophthalmoplegia	
7445 U → C	$tRNA^{Ser(UCN)}$	Non-syndromic deafness	Levinger et al. (2001)
3243 A → G[c]	$tRNA^{Leu(UUR)}$	MELAS	Levinger et al. (2004)
3302 A → G		Myopathy	
3303 C → T		Cardiomyopathy	
Hypothesized processing defects			
4469 C → A	$tRNA^{Met}$	Maternally inherited hypertension	Liu et al. (2017)

[a]Based on the position in the mitochondrial genome (Anderson et al. 1981)
[b]Aberrant structure as well
[c]Reduced aminoacylation also (Park et al. 2003)

there is evidence that mutations in various regions of the mt-tRNA can reduce the level of 3′-end processing. This may be due to the different mechanisms by which mtRNase P and RNase Z interact with their substrates. As mentioned above, mtRNase P appears to cleave mt-tRNAs first from mt-RNAs and thus may require less structural information from the folded mt-tRNAs compared to RNase Z. Indeed, nucleus-encoded tRNAs and mt-tRNAs can form alternative structures before base modification occurs, and base modifications are required for normal tertiary structure. The initial extended hairpin structures that form after transcription may be efficiently recognized by RNase P, but not RNase Z. Also, MRPP2/TRMT10C is responsible for the m^1A9 methylation in mitochondria, which has been shown to be important for tertiary folding for some mt-tRNAs. Therefore, this modification by the MRRP1/MRPP2 subcomplex may be required before RNase Z is able to recognize the mt-tRNA. Adding further complexity, mt-tRNA ends are in different contexts. The secondary structure each mt-tRNA assumes after transcription presumably affects the directly neighboring mt-RNA. For example, $mt\text{-}tRNA^{Leu(UUR)}$ is between mt-rRNA and ND1, whereas $mt\text{-}tRNA^{Ser(AGY)}$ is directly between two mt-tRNAs (Fig. 3.1).

Defects in 5′-end processing have been identified in patients suffering from maternally inherited hypertension (Guo et al. 2016; Jiang et al. 2016; Li et al. 2009; Wang et al. 2011; Zhu et al. 2009). Since mitochondria are loaded into the oocyte, this non-Mendelian form of inheritance is the first indication that the phenotype is due to mutations in mtDNA. Numerous mtDNA mutations are now associated with maternally inherited hypertension. Wang et al. were the first to identify a novel mtDNA mutation 4263A → G that altered 5′-end mt-tRNA processing (Wang et al. 2011). The causes of hypertension are not well understood but are thought to arise from complex changes in different tissues (Page 1967). However, the production of reactive oxygen species (ROS) and inflammation appear to be two underlying causes, and damaged mitochondria are in general frequently the source of abnormal levels of ROS. Wang et al. identified homoplasmic females harboring the 4263A → G mutation, which occurs in position 1 of mt-tRNA^{Ile} (Fig. 3.4b, Wang et al. 2011). This mutation results in decreased levels of mt-tRNA^{Ile}, mt-RNA translation, and respiration. Using an in vitro mtRNase P reconstitution assay, the authors demonstrated that the 4263A → G mutation decreased 5′-end processing efficiency by ~30% compared to wild type. Recently, Jiang et al. identified an additional mutation in mt-tRNA^{Ala} associated with hypertension in three unrelated individuals whose families exhibited maternally inherited hypertension (Jiang et al. 2016). The 5655A → G mutation is in the same relative position in the tRNA as the 4263A → G mutation and causes the same mitochondrial deficits, such as reduced mt-tRNA^{Ala} levels, decreased respiration, and a 35% decrease in 5′-end processing (Fig. 3.4f). In addition, cybrids derived from patient lymphocytes exhibited reduced oxygen consumption and ATP levels, increased ROS, and improperly aminoacylated mt-tRNA^{Ala} (Jiang et al. 2016).

Three reported mt-tRNA mutations in the literature have not been directly tested for loss of 5′-end processing, but they are in a location that strongly suggests a causal effect on a processing defect. Two groups independently identified 4401A → G in maternally hypertensive patients (Li et al. 2009; Zhu et al. 2009). This mutation is particularly interesting as it affects both mt-tRNA^{Gln} (encoded on the L-strand) and mt-tRNA^{Met} (encoded on the H-strand) at their RNase P cleavage sites (Fig. 3.4c, d). Using patient-derived cybrid cells, both groups showed decreased mt-tRNA levels and respiration, and Li et al. also showed a decrease in mitochondrial protein synthesis. More recently, Guo et al. identified a 5512A → G mutation in mt-tRNA^{Trp} in a family of maternally inherited hypertensives (Guo et al. 2016). This mutation is at the analogous position to the 4263A → G mutation in mt-tRNA^{Ile} and thus likely also affects 5′-end processing (Fig. 3.4e).

The investigators who uncovered mtDNA mutations that affect 5′-end processing were interested in causes of maternally inherited hypertension. To identify mt-tRNA mutations affecting 3′-end processing, Levinger et al. instead started with mt-tRNA mutations known to cause mitochondrial diseases and syndromes (Levinger et al. 2001, 2003, 2004). The U7445 U → C mutation in mt-tRNA$^{Ser(UCN)}$ was known to cause non-syndromic deafness, reduced mt-tRNA $^{Ser(UCN)}$ levels and mitochondrial protein synthesis (Guan et al. 1998; Reid et al. 1997). This mutation is located precisely at the

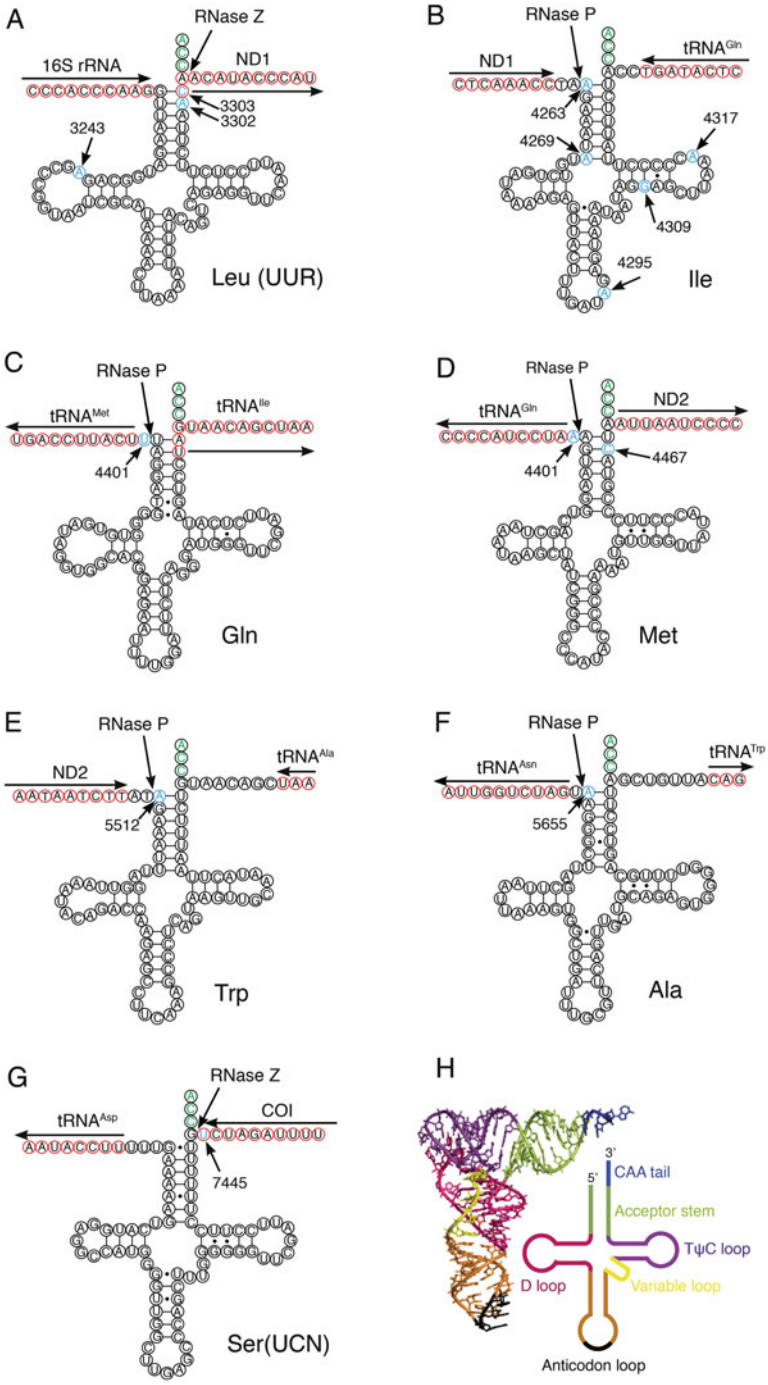

Fig. 3.4 Human pathogenic mt-tRNA mutations. (**a**–**g**) Human mt-tRNA cloverleaf structures encoded from left to right of human mtDNA (Fig. 3.1a). Blue nucleotides represent pathogenic

RNase Z cleavage site (Fig.3.4g). Using in vitro-labeled mt-tRNA $^{Ser(UCN)}$ and mitoplast extract from HeLa cells as a source of RNase Z, Levinger et al. found the mutant mt-tRNA was unable to undergo any 3′-end processing in contrast to wild-type mt-tRNA even though they both appeared to form normal structures (Levinger et al. 2001). To extend these studies, the authors examined the effect on 3′-end processing of mutations in mt-tRNA^{Ile} and mt-tRNA$^{Leu(UUR)}$, two mt-tRNAs linked to cardiomyopathies, ophthalmoplegia, and mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes (MELAS) (Levinger et al. 2003, 2004). Using experiments similar to their previous work, the authors found four mutations in mt-tRNA^{Ile} and one in mt-tRNA$^{Leu(UUR)}$ that substantially reduced 3′-end processing (Fig. 3.4a, b). Some of these mutations also resulted in structural and aminoacylation changes (Table 3.1).

An additional mt-tRNA mutation was recently published that was not analyzed for defective 3′-end processing but that may interfere based on its position. Liu et al. identified the 4467C → A mutation in mt-tRNA^{Met} in a family with maternally inherited hypertension (Liu et al. 2017). This mutation is located near the RNase Z cleavage site (Fig. 3.4d). Using Lymphocyte cell lines derived from patients, Liu et al. demonstrated decreased ATP in cell extract, increased oxidative damage, and lower oxygen consumption, indicating disrupted mitochondrial function.

3.4.2 Mutations in Mitochondrial RNase P

Human diseases due to mitochondrial deficiencies have now been linked to mutations in all three members of the mtRNase P complex (Table 3.2). They are rare diseases and many of the mutations occur in highly conserved residues. MRPP2/HSD10 (encoded by the *HSD17B10* gene) has the largest number of identified mutations, referred to as HSD10, and was the first to be described [OMIM #300438, 300,256 (reviewed in Zschocke 2012)]. For consistency, we will refer to the protein as MRPP2 in this section; however, HSD10 disease is the name of diseases caused by mutations in *HSD17B10*. As described above, MRPP2 is a multifunctional protein. It functions in isoleucine metabolism but has been shown to be a promiscuous dehydrogenase in vitro (reviewed in Moeller and Adamski 2009). Because MRPP2 was identified as a member of the mtRNase P complex and plays a role in mt-tRNA 5′-end processing,

Fig. 3.4 (continued) mutations (Table 3.1). The numbering system is according to (Anderson et al. 1981). Arrows at the 5′- and 3′-ends indicate the mtRNase P and RNase Z sites, respectively, impacted by mutation. Red nucleotides indicate the neighboring 5′ and 3′ transcripts. The arrows above and below the red nucleotides indicate the start (beginning of arrow) or stop (arrowhead) of the transcript close to the cleavage sites. The flanking gene may overlap with the mt-tRNA (e.g., mt-tRNA$^{Leu(UUR)}$, **a**) or there may be intervening nucleotides (e.g., mt-tRNA^{Trp}, **e**). The conserved CCA sequence (green) found on the acceptor stem at the 3′-end of all tRNAs acts as the amino acid attachment site and is uncoded; tRNA-nucleotidyltransferase is responsible for CCA addition. (**h**) Tertiary structure of yeast tRNA^{Phe} color-coded for the D loop, anticodon loop, variable loop, TψC loop, acceptor stem, and CAA tail as indicated by the cloverleaf cartoon

Table 3.2 Disease phenotypes due to mitochondrial RNase P complex mutations

	Human disease/ phenotype	Citation(s)[a]
MRPP2/HSD10 (Scully)		
D86G	Neonatal onset	Rauschenberger et al. (2010)
R226Q		Perez-Cerda et al. (2005) and Vilardo and Rossmanith (2015)
N247S		Chatfield et al. (2015) and Vilardo and Rossmanith (2015)
V12 L	Infantile onset	Oerum et al. (2017)
L122 V		Ofman et al. (2003)
R130C		Deutschmann et al. (2014) and Vilardo and Rossmanith (2015)
P210S		Vilardo and Rossmanith (2015)
V176 M		Oerum et al. (2017)
E249Q	Juvenile	Yang et al. (2009)
K212E		Falk et al. (2016)
V65A		Richardson et al. (2016) and Seaver et al. (2011)
Q165H	Atypical/adult onset/ nonregressive	Rauschenberger et al. (2010)
A154T		Fukao et al. (2014)
A157V		Akagawa et al. (2017)
MRPP1/TRMT10C (Rswl)		
R181L	Infantile lethality	Metodiev et al. (2016)
T272A		
MRPP3 (Mldr)		
A485V	Perrault syndrome	Hochberg et al. (2017)
ELAC2 (dRNase Z^L)		
F154 L	Variable—Infantile to juvenile	Haack et al. (2013)
Arg211Stop, T520I		
L423F		
chr17:12,903,471A → T (splice site mutation)		Akawi et al. (2016)

[a]The most relevant citation(s) for mt-RNA processing. First reported cases and symptoms summarized in Akagawa et al. (2017) and Zschocke (2012)

ascribing its molecular role in disease has been complicated (Holzmann et al. 2008). However, the primary cause of symptoms experienced by patients with MRPP2 mutations may be due to loss of mitochondrial function and mt-RNA processing (Zschocke 2012). For example, no clinical correlation has been found between loss of the dehydrogenase activity and symptoms. The mutation Q165H, which abolishes this activity, does not cause the classical, more severe disease onset, and those mutations that are not near the enzyme's active center cause the neonatal form of the disease (Rauschenberger et al. 2010). In addition, reduced dietary isoleucine does not improve HSD10 symptoms (Korman 2006; Sutton et al. 2003; Zschocke et al. 2000). In vitro work has also shown dehydrogenase-dead protein can still rescue vital cellular

functions and functions in the mtRNase P complex (Rauschenberger et al. 2010; Vilardo et al. 2012).

HSD10 disease is rare with no effective treatment and exhibits variable ages of onset. Symptoms are usually multisystemic, consistent with mitochondrial disease. Patients often have a progressive loss of cognitive and motor functions, with epilepsy and blindness. The most severe forms result in cardiomyopathy and severe neurodegeneration (Zschocke 2012). Mutations can affect homotetramization, the dehydrogenase activity or its ability to function in the mtRNase P complex (Vilardo and Rossmanith 2015). As *HSD17B10* is X-linked, mutations mostly affect males, but heterozygous females can also have disease symptoms, making the inheritance pattern X-linked dominant.

After MRPP2 was identified as a member of the mtRNase P complex, researchers began testing samples from patients with HSD10 disease, as well as recombinant pathogenic MRPP2 mutations, for mitochondrial deficits and decreased mt-RNA processing, along with dehydrogenase and methyltransferase activities (Chatfield et al. 2015; Deutschmann et al. 2014; Falk et al. 2016; Oerum et al. 2017; Vilardo and Rossmanith 2015). Deutschmann et al. first showed that patient fibroblast cells harboring R130C, the most prevalent mutation, significantly accumulate unprocessed mt-RNAs encoded on the mtDNA heavy strand, but not the light strand (Deutschmann et al. 2014). Chatfield et al. examined post-autopsy tissues from a patient with the N247S mutation and found increased unprocessed mt-RNAs and decreased mitochondrial protein synthesis (Chatfield et al. 2015). Vilardo and Rossmanith took an in vitro approach, examining four different mutations, R226Q, N247S, P210S, and R130C. All four mutations greatly impair dehydrogenase activity, mt-tRNA processing, and methylation. R226Q and N247S mutations also disrupted homotetramer formation and interaction with MRPP1 (Vilardo and Rossmanith 2015). The MRPP1/2 subcomplex is highly stable. There is evidence that reducing the amount of MRPP2 can affect the stability of MRPP1. Deutschmann et al. examined protein levels in patient fibroblasts and found the reduced level of MRPP2 resulted in lower MRPP1 levels, but not MRPP3 levels, whereas loss of MRPP1 had no effect on MRPP2 levels (Deutschmann et al. 2014). In addition, Falk et al. identified a novel K212E mutation and showed that while it has only a modest effect on dehydrogenase activity, it reduced methylation and 5′-end processing activities, as well as binding to MRPP1 (Falk et al. 2016). Most recently, Oerum et al. identified the novel mutations V12 L and V176 M. Both reduced dehydrogenase, methyltransferase, and mtRNase P activities, with V176 M being much more severe, in agreement with patient symptoms (Oerum et al. 2017).

As with MRPP2, the dual function of MRPP1 as a methyltransferase and member of the mtRNase P complex makes it challenging to ascribe the primary defect of mitochondrial loss of function. Disease-causing mutations in MRPP1/TRMT10C have recently been identified (Metodiev et al. 2016). As mentioned above, the MRPP1/2 subcomplex is an active methyltransferase that can methylate adenosine and guanine nucleotides at position 9 (Vilardo et al. 2012). Nineteen out of the 22 mt-tRNAs have an A or G at position 9, and thus, they are likely acted upon by MRPP2. Metodiev et al. described two unrelated patients suffering from lactic

acidosis, hypotonia, feeding difficulties, and deafness. The patients' symptoms were so severe that they died after only 5 months of age. After whole-exome sequencing, they were found to harbor mutations in MRPP1 (Table 3.2). Both patients exhibited clear signs of mitochondrial disease, with deficits in Complexes I and IV. MRPP1 protein levels were reduced in patient-derived fibroblast cell lines, while MRPP2 and MRPP3 levels were unchanged. The amount of m^1R9 methyltransferase activity was also unchanged. The steady-state levels of mt-tRNAs and mt-mRNAs remained mostly unchanged in patient-derived fibroblasts; however there was an increase in unprocessed mt-RNAs that was rescued by transfecting the cells with wild-type MRPP1. These results led the authors to speculate that the MRPP1 mutations may affect interaction with MRPP3 rather than MRPP2 and the increased amount of unprocessed mt-RNAs may interfere with translation.

Hochberg et al. have identified a disease-causing mutation in MRPP3 (Hochberg et al. 2017). Perrault syndrome is a rare, genetically heterogeneous disease characterized by sensorineural hearing loss in males and females and primary ovarian insufficiency (Jenkinson et al. 2012). Most of the genes mutated in Perrault syndrome are involved in mitochondrial translation (OMIM #233400). Hochberg et al. identified a family affected by Perrault syndrome of which three individuals harbored a A485V mutation in MRPP3. Located in the metallonuclease domain, A485V does not cause reduced protein levels in patient-derived fibroblasts but does decrease the level of mtDNA-encoded respiratory chain complex proteins. In addition, the authors detected multiple unprocessed mt-RNAs. The authors went on to test the A485V mutation in an in vitro mtRNase P reconstitution assay and found the mutated MRPP3 possessed significantly less 5′-end processing activity compared to wild type.

Since mutations in mtRNase P clearly affect mt-RNA processing, causing mitochondrial dysfunction and disease, it is possible that defects in 3′-end mt-RNA processing could do the same. Haack et al. performed whole-exome sequencing on patients suffering from oxidative phosphorylation deficiencies and identified four disease-associated alleles of ELAC2 in three families (Table 3.2, Haack et al. 2013). Using qPCR, they found an increase in unprocessed mt-RNA in patient tissue samples. The processing defect could be rescued in patient fibroblasts by lentivirus-mediated infection with wild-type ELAC2. However, mt-tRNA levels and at least four mt-mRNAs were unchanged, and there was no evidence of 3′-end processing defects. Despite this, mitochondrial translation and protein levels were reduced, suggesting that the increase in unprocessed mt-RNA intermediates may hamper mitochondrial translation. Akawi et al. recently identified a splicing mutation that reduces ELAC2 levels (Akawi et al. 2016). Using real-time PCR, they found increased levels of unprocessed mt-RNAs in fibroblasts of one patient.

3.5 Loss of Mt-tRNA Processing in Model Organisms

Model organisms are powerful systems for studying the cellular and molecular bases underlying human disease. Human mutations in MRPP1, MRPP2, and MRPP3 are rarely null alleles. This is because complete absence of mtRNase P activity is likely incompatible with life. Because we can carry out conditional knockouts and rigorous genetics, model organisms offer the opportunity to study the in vivo effects of complete loss of mtRNase P activity on mt-RNA processing and mitochondrial function, as well as the dehydrogenase and methyltransferase activities of MRPP1 and MRPP2. Single orthologs for each of the three mtRNase P complex members have been identified in mouse and *Drosophila*. MRPP2 is the smallest protein and has very high amino acid identity among human, mouse, and *Drosophila* (Fig. 3.2). The crystal structure of human HSD10 was determined as a homotetramer complexed with NAD^+ and an inhibitor (Kissinger et al. 2004). Given the high amino acid identity, this structure was used to model the structure for the *Drosophila* MRPP2 Scully (Fig. 3.2d, e). MRPP3 contains seven pentatricopeptide repeats and a metallonuclease domain (Fig. 3.2; Howard et al. 2012). There is only one MRPP3 in all three species, with the highest identity in the metallonuclease domain (Fig. 3.2). MRPP1/TRMT10C also has only one ortholog in humans, mouse, and *Drosophila*. Shao et al. solved the structure for yeast TRM10 in the presence and absence of a methyl donor and showed that the catalytic domain displays the typical SpoU-TrmD fold found in SPOUT family methyltransferases (Fig. 3.2; Shao et al. 2014).

3.5.1 Drosophila Mitochondrial RNase P

Drosophila is the only model organism in which all three members of the mtRNase P complex have been studied and in which MRPP3's essential role in vivo was first demonstrated (Table 3.3; Sen et al. 2016). Sen et al. identified the MRPP3 and MRPP1 orthologs, called Mulder (Mldr) and Roswell (Rswl), respectively, and used genetics and cell biology to determine the role of mtRNase P in mitochondrial function and mt-RNA processing in vivo. All three *Drosophila* orthologs localize to mitochondria and associate with each other (Sen et al. 2016). Torroja et al. characterized mutations in *scu* and found they were lethal and had reduced dehydrogenase activity (Torroja et al. 1998). By examining EMS-induced point mutations, as well as RNAi knockdown, Sen et al. showed that loss of each mtRNase P complex member delayed larval development and was pupal-lethal. Tissue extract from mutants had very low ATP levels, and immunofluorescence showed swollen mitochondria. Using Northern blots probed with four different mt-tRNAs in different RNA contexts, the authors found unprocessed mt-RNAs in extract from *rswl* RNAi knockdown and *mldr* and *scu* mutants. While Sen et al. only had RNAi available for *rswl*, we can show that a transposable element-induced mutation in

Table 3.3 Model organism phenotypes due to loss of mtRNase P complex

	Phenotype	Citation(s)
Scully (MRPP2/HSD10)		
Mouse		
Knockout	Embryonic lethality	Rauschenberger et al. (2010)
Conditional knockout	(endothelial cells) spleen, vasculature abnormalities, die at 25 weeks (noradrenergic neurons) die 26 weeks	
Xenopus		
Antisense knockdown		Rauschenberger et al. (2010)
Drosophila		
S163F (S169F) *scu*^{D}	Pupal lethal	Sen et al. (2016)
Q159Stop *scu*^{A}		
E205X *scu*^{4058}		Sen et al. (2016) and Torroja et al. 1998
L33Q *scu*^{174}		Torroja et al. (1998)
A86 + 7Stop *scu*^{3127}		
scu$^{TriP.HMS02305}$ *RNAi*		Sen et al. (2016)
scu$^{TriP.GL01079}$ *RNAi*		
Roswell (MRPP1)		
Drosophila		
rswl^{07838}	Pupal lethal	This work
rswl^{GD12447} RNAi		Sen et al. (2016)
rswl$^{TriP.HMC02423}$ RNAi		
Mulder (MRPP3)		
Mouse		
Knockout	Embryonic lethality (heart/ skeletal muscle) cardiomyopathy	Rackham et al. (2016)
Conditional knockout		
Drosophila		
W465A (W520A) *mldr*^{C}	Pupal lethal	Sen et al. (2016)
Y121D Y183D *mldr*^{B}		
mldr^{KK108043} RNAi		
dRNase Z (ELAC2)		
Drosophila		
dRNaseZ^{v43751} RNAi	Pupal lethal	Xie et al. (2011)
dRNaseZ^{v43521} RNAi		
dRNaseZ^{ED24}		Xie et al. (2013)

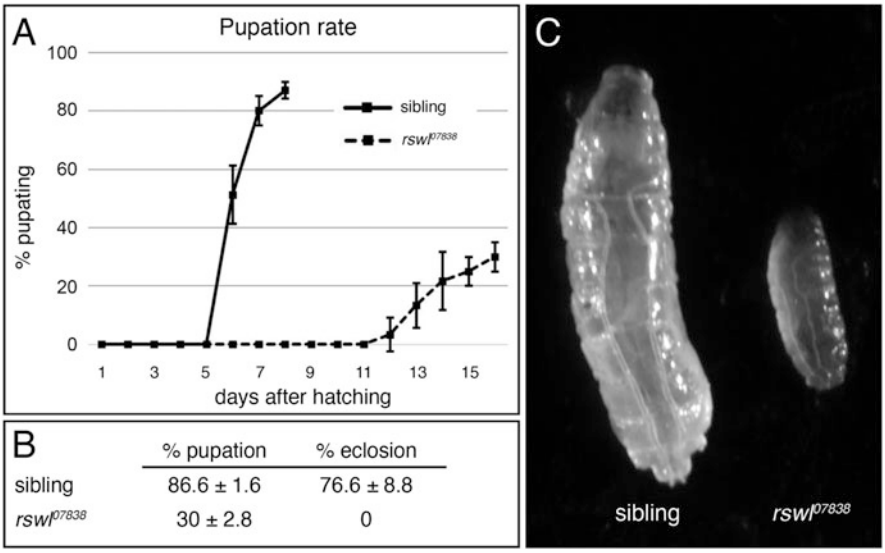

Fig. 3.5 A mutation in *roswell* causes pupal lethality. (**a**, **b**) A graph showing *rswl*^{07838} mutant larvae have delayed development and only approximately 30% eventually pupate compared to sibling controls. No *rswl*^{07838} mutant adults emerge from the pupal cases (eclosion) (**b**). *rswl*^{07838} is induced by a transposable element insertion (Bellen et al. 2004; Spradling et al. 1999). (**c**) After 5 days, *rswl*^{07838} mutant larvae (left) are much smaller than their wild-type siblings (right). These phenotypes are consistent with *rswl* RNAi knockdown and loss of *Mulder* and *Scully* (Sen et al. 2016). Pupation and eclosion rates were performed as described in Sen et al. (2016)

rswl is also lethal, with delayed pupal development consistent with our previous observations (Fig. 3.5, M.S. and R.T.C., unpublished data).

Mouse and human contain two proteins, ELAC1 and ELAC2, that perform 3′-end processing of nuclear and mitochondrial tRNAs, respectively (see above). ELAC2 encodes two products of which only one is targeted to mitochondria. In contrast, *Drosophila* contains a single gene, dRNaseZ, which encodes an N-terminal mitochondrial targeting sequence and two nuclear localization signals (NLS). It is not clear how the protein gets differentially targeted to the two organelles, though there are other examples of this situation in the literature (Yogev and Pines 2011). Dubrovsky et al. cloned the *Drosophila* homolog and showed that it can cleave tRNAs in vitro (Dubrovsky et al. 2004). They also demonstrated that RNAi knockdown and deletion of the gene causes growth defects and defects in tRNA and mt-tRNA processing (Table 3.3; Xie et al. 2011, 2013). However, it is not clear whether the lethality due to loss of the gene is caused by mitochondria-specific disruptions, nucleus disruptions, or both.

3.5.2 Mouse Mitochondrial RNase P

Rackham et al. created knockout mice for MRPP3 (Table 3.3; Rackham et al. 2016). The authors not only clearly showed mt-tRNA processing defects but were also able to show that mt-RNA processing links transcription to translation through mitoribosome assembly (described above). Using Cre recombinase, the authors removed the third exon of MRPP3 to create a knockout allele. The MRPP3 full-body conditional knockout mice died at day E8.5. This is consistent with mouse knockouts for proteins involved in mitochondrial gene expression. Since the mice died so young, Rackham et al. used muscle-specific Cre to produce mice lacking MRPP3 in heart and skeletal muscle. The mice had reduced muscle fibers with reduced Complex I and Complex IV staining, and died at week 11 from cardiomyopathy, a phenotype frequently seen in mitochondrial disease.

MRPP2/HSD10 is the other member of the mtRNase P complex that has been studied in mouse. Rauschenberger et al. created a conditional knockout of MRPP2 (Table 3.3; Rauschenberger et al. 2010). They reported that the mouse knockout resulted in early embryonic lethality. In order to study MRPP2 function, they established conditional knockouts, one using a Cre recombinase to eliminate the protein in endothelial and immune cells and one using Cre to affect noradrenergic neurons. The endothelial knockout mice died at week 25 with defects in spleen and vasculature. The noradrenergic knockout mice died at week 26. Mitochondria in the loci coerulei of the brain, which contain noradrenergic neurons, lacked normal, dense cristae. To circumvent the early embryonic lethality in mice, Rauschenberger et al. also examined MRPP2 function in another vertebrate, *Xenopus laevis*, using morpholinos to knock down protein expression in the animal cap (Rauschenberger et al. 2010). Explants from these cells had decreased mitochondrial function inferred by a decrease in pyruvate turnover. Mitochondrial ultrastructure was also disrupted as judged by transmission electron microscopy. In addition, the authors examined morpholino-induced loss of MRPP2 on neural tissue patterning. They found patterning was not disrupted, but tissue size was reduced. This may have been due to an increase in apoptosis, as shown with increased TUNEL labeling.

3.6 Concluding Remarks and Future Directions

Processing of mtDNA transcripts presents unique challenges due to the constraints of the very small genome encoding all three major RNA types encoded in polycistronic messages. mtDNA supplies the mt-rRNAs for the mitoribosome and all the mt-tRNAs. Both types of mt-RNA are required in order to translate the 13 proteins encoded by mtDNA. As such, properly cleaving each product becomes essential in order for the organelle to maintain its function. The suite of mt-tRNAs is also small relative to the number encoded in the nucleus. There is only one cognate mt-tRNA

per amino acid, with the exception of serine and leucine; thus, disrupting processing of a single mt-tRNA transcript is problematic.

We now know that point mutations in mt-tRNAs can affect processing, leading to maternally inherited hypertension and mitochondrial disease. In addition, pathogenic mutations in MRPP1, MRPP2, and MRPP3 have been identified. Because null alleles of mtRNase P are likely to be lethal, they have not appeared in humans. Given the diverse nature of symptoms resulting from defects in mt-tRNA processing, especially the link to hypertension, there may be many more as yet undetected mild mutations that could predispose individuals to a variety of skeletal, neural, and cardiovascular problems.

This is an exciting time to study mt-tRNA processing. Mutations in mt-tRNAs have long been known to cause mitochondrial diseases, but we are still parsing the details of how mtDNA transcription, processing, and translation control feedback to each other, the sequential steps of regulation, and how defects in each process lead to different diseases and symptoms. A better understanding of the molecular mechanisms underlying mt-RNA processing may open avenues to more effective treatments and cures of those with mitochondrial diseases.

Funding This work was supported by the National Institutes of Health/Department of Defense [CHIRP HU0001–14–2-0041 to M.S. and R.T.C.].

References

Akagawa S et al (2017) Japanese male siblings with 2-methyl-3-hydroxybutyryl-CoA dehydrogenase deficiency (HSD10 disease) without neurological regression. JIMD Rep 32:81–85. https://doi.org/10.1007/8904_2016_570

Akawi NA et al (2016) A homozygous splicing mutation in *ELAC2* suggests phenotypic variability including intellectual disability with minimal cardiac involvement. Orphanet J Rare Dis 11:139. https://doi.org/10.1186/s13023-016-0526-8

Anderson S et al (1981) Sequence and organization of the human mitochondrial genome. Nature 290:457–465

Antonicka H, Shoubridge EA (2015) Mitochondrial RNA granules are centers for posttranscriptional RNA processing and ribosome biogenesis. Cell Rep. https://doi.org/10.1016/j.celrep.2015.01.030

Antonicka H, Sasarman F, Nishimura T, Paupe V, Shoubridge EA (2013) The mitochondrial RNA-binding protein GRSF1 localizes to RNA granules and is required for posttranscriptional mitochondrial gene expression. Cell Metab 17:386–398. https://doi.org/10.1016/j.cmet.2013.02.006

Bellen HJ et al (2004) The BDGP gene disruption project: single transposon insertions associated with 40% of Drosophila genes. Genetics 167:761–781. https://doi.org/10.1534/genetics.104.026427

Borowski LS, Dziembowski A, Hejnowicz MS, Stepien PP, Szczesny RJ (2013) Human mitochondrial RNA decay mediated by PNPase–hSuv3 complex takes place in distinct foci. Nucleic Acids Res 41:1223–1240. https://doi.org/10.1093/nar/gks1130

Brandon MC, Lott MT, Nguyen KC, Spolim S, Navathe SB, Baldi P, Wallace DC (2005) MITOMAP: a human mitochondrial genome database—2004 update. Nucleic Acids Res 33:D611–D613. https://doi.org/10.1093/nar/gki079

Brzezniak LK, Bijata M, Szczesny RJ, Stepien PP (2011) Involvement of human ELAC2 gene product in 3′ end processing of mitochondrial tRNAs. RNA Biol 8:616–626. https://doi.org/10.4161/rna.8.4.15393

Chatfield KC et al (2015) Mitochondrial energy failure in HSD10 disease is due to defective mtDNA transcript processing. Mitochondrion 21:1–10. https://doi.org/10.1016/j.mito.2014.12.005

Chinnery PF (2000) Mitochondrial disorders overview. In: Adam MP, Ardinger HH, Pagon RA et al (eds) GeneReviews® [Internet]. University of Washington, Seattle, pp 1993–2017

Claros MG, Vincens P (1996) Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur J Biochem 241:779–786

Deutschmann AJ et al (2014) Mutation or knock-down of 17beta-hydroxysteroid dehydrogenase type 10 cause loss of MRPP1 and impaired processing of mitochondrial heavy strand transcripts. Hum Mol Genet 23:3618–3628. https://doi.org/10.1093/hmg/ddu072

Dubrovsky EB, Dubrovskaya VA, Levinger L, Schiffer S, Marchfelder A (2004) *Drosophila* RNase Z processes mitochondrial and nuclear pre-tRNA 3′ ends *in vivo*. Nucleic Acids Res 32:255–262. https://doi.org/10.1093/nar/gkh182

Falk MJ et al (2016) A novel *HSD17B10* mutation impairing the activities of the mitochondrial RNase P complex causes X-linked intractable epilepsy and neurodevelopmental regression. RNA Biol 13:477–485. https://doi.org/10.1080/15476286.2016.1159381

Fukao T et al (2014) The first case in Asia of 2-methyl-3-hydroxybutyryl-CoA dehydrogenase deficiency (HSD10 disease) with atypical presentation. J Hum Genet 59:609–614. https://doi.org/10.1038/jhg.2014.79

Gobert A et al (2010) A single *Arabidopsis* organellar protein has RNase P activity. Nat Struct Mol Biol 17:740. doi:https://doi.org/10.1038/nsmb.1812. https://www.nature.com/articles/nsmb.1812—supplementary-information

Gray MW (2012) Mitochondrial evolution. Cold Spring Harb Perspect Biol 4:a011403. https://doi.org/10.1101/cshperspect.a011403

Guan MX, Enriquez JA, Fischel-Ghodsian N, Puranam RS, Lin CP, Maw MA, Attardi G (1998) The deafness-associated mitochondrial DNA mutation at position 7445, which affects tRNA$^{Ser(UCN)}$ precursor processing, has long-range effects on NADH dehydrogenase subunit ND6 gene expression. Mol Cell Biol 18:5868–5879

Guo L, Yuan Y, Bi R (2016) Mitochondrial DNA mutation m.5512A > G in the acceptor-stem of mitochondrial tRNA^{Trp} causing maternally inherited essential hypertension. Biochem Biophys Res Commun 479:800–807. https://doi.org/10.1016/j.bbrc.2016.09.129

Haack TB et al (2013) ELAC2 mutations cause a mitochondrial RNA processing defect associated with hypertrophic cardiomyopathy. Am J Hum Genet 93:211–223. https://doi.org/10.1016/j.ajhg.2013.06.006

Hochberg I et al (2017) A homozygous variant in mitochondrial RNase P subunit PRORP is associated with Perrault syndrome characterized by hearing loss and primary ovarian insufficiency. BioRxiv. https://doi.org/10.1101/168252

Holzmann J, Frank P, Löffler E, Bennett KL, Gerner C, Rossmanith W (2008) RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. Cell 135:462–474. https://doi.org/10.1016/j.cell.2008.09.013

Howard MJ, Lim WH, Fierke CA, Koutmos M (2012) Mitochondrial ribonuclease P structure provides insight into the evolution of catalytic strategies for precursor-tRNA 5′ processing. Proc Natl Acad Sci U S A 109:16149–16154. https://doi.org/10.1073/pnas.1209062109

Iborra FJ, Kimura H, Cook PR (2004) The functional organization of mitochondrial genomes in human cells. BMC Biol 2:9. https://doi.org/10.1186/1741-7007-2-9

Jablonski JA, Caputi M (2009) Role of cellular RNA processing factors in human immunodeficiency virus type 1 mRNA metabolism, replication, and infectivity. J Virol 83:981–992. https://doi.org/10.1128/JVI.01801-08

Jackman JE, Montange RK, Malik HS, Phizicky EM (2003) Identification of the yeast gene encoding the tRNA m^1G methyltransferase responsible for modification at position 9. RNA 9:574–585. https://doi.org/10.1261/rna.5070303

Jenkinson EM et al (2012) Perrault syndrome: further evidence for genetic heterogeneity. J Neurol 259:974–976. https://doi.org/10.1007/s00415-011-6285-5

Jiang P et al (2016) A hypertension-associated $tRNA^{Ala}$ mutation alters tRNA metabolism and mitochondrial function. Mol Cell Biol 36:1920–1930. https://doi.org/10.1128/MCB.00199-16

Jourdain AA, Koppen M, Wydro M, Rodley CD, Lightowlers RN, Chrzanowska-Lightowlers ZM, Martinou JC (2013) GRSF1 regulates RNA processing in mitochondrial RNA granules. Cell Metab 17:399–410. https://doi.org/10.1016/j.cmet.2013.02.005

Kash JC, Cunningham DM, Smit MW, Park Y, Fritz D, Wilusz J, Katze MG (2002) Selective translation of eukaryotic mRNAs: functional molecular analysis of GRSF-1, a positive regulator of influenza virus protein synthesis. J Virol 76:10417–10426. https://doi.org/10.1128/JVI.76.20.10417-10426.2002

Kissinger CR et al (2004) Crystal structure of human ABAD/HSD10 with a bound inhibitor: implications for design of Alzheimer's disease therapeutics. J Mol Biol 342:943–952. https://doi.org/10.1016/j.jmb.2004.07.071

Klemm BP, Wu N, Chen Y, Liu X, Kaitany KJ, Howard MJ, Fierke CA (2016) The diversity of ribonuclease P: protein and RNA catalysts with analogous biological functions. Biomol Ther 6. https://doi.org/10.3390/biom6020027

Korman SH (2006) Inborn errors of isoleucine degradation: a review. Mol Genet Metab 89:289–299. https://doi.org/10.1016/j.ymgme.2006.07.010

Land M et al (2015) Insights from 20 years of bacterial genome sequencing. Funct Integr Genomics 15:141–161. https://doi.org/10.1007/s10142-015-0433-4

Levinger L, Jacobs O, James M (2001) *In vitro* 3′-end endonucleolytic processing defect in a human mitochondrial $tRNA^{Ser(UCN)}$ precursor with the U7445C substitution, which causes non-syndromic deafness. Nucleic Acids Res 29:4334–4340

Levinger L, Giege R, Florentz C (2003) Pathology-related substitutions in human mitochondrial $tRNA^{Ile}$ reduce precursor 3′ end processing efficiency *in vitro*. Nucleic Acids Res 31:1904–1912

Levinger L, Oestreich I, Florentz C, Mörl M (2004) A pathogenesis-associated mutation in human mitochondrial $tRNA^{Leu(UUR)}$ leads to reduced 3′-end processing and CCA addition. J Mol Biol 337:535–544. https://doi.org/10.1016/j.jmb.2004.02.008

Lewis OL, Farr CL, Kaguni LS (1995) *Drosophila melanogaster* Mitochondrial DNA: completion of the nucleotide sequence and evolutionary comparisons. Insect Mol Biol 4:263–278. https://doi.org/10.1111/j.1365-2583.1995.tb00032.x

Li R, Liu Y, Li Z, Yang L, Wang S, Guan MX (2009) Failures in mitochondrial $tRNA^{Met}$ and $tRNA^{Gln}$ metabolism caused by the novel 4401A>G mutation are involved in essential hypertension in a Han Chinese family. Hypertension 54:329–337. https://doi.org/10.1161/HYPERTENSIONAHA.109.129270

Liu Y, Li Y, Zhu C, Tian L, Guan M, Chen Y (2017) Mitochondrial biogenesis dysfunction and metabolic dysfunction from a novel mitochondrial $tRNA^{Met}$ 4467 C>A mutation in a Han Chinese family with maternally inherited hypertension. Sci Rep 7:3034. https://doi.org/10.1038/s41598-017-03303-w

Margulis L (1970) Origin of eukaryotic cells. Yale University Press, New Haven

Mercer TR et al (2011) The human mitochondrial transcriptome. Cell 146:645–658. https://doi.org/10.1016/j.cell.2011.06.051

Metodiev MD et al (2016) Recessive mutations in *TRMT10C* cause defects in mitochondrial RNA processing and multiple respiratory chain deficiencies. Am J Hum Genet 98:993–1000. https://doi.org/10.1016/j.ajhg.2016.03.010

Moeller G, Adamski J (2009) Integrated view on 17beta-hydroxysteroid dehydrogenases. Mol Cell Endocrinol 301:7–19. https://doi.org/10.1016/j.mce.2008.10.040

Oerum S et al (2017) Novel patient missense mutations in the HSD17B10 gene affect dehydrogenase and mitochondrial tRNA modification functions of the encoded protein. Biochim Biophys Acta 1863(12):3294–3302. https://doi.org/10.1016/j.bbadis.2017.09.002

Ofman R et al (2003) 2-Methyl-3-hydroxybutyryl-CoA dehydrogenase deficiency is caused by mutations in the HADH2 gene. Am J Hum Genet 72:1300–1307

Ojala D, Montoya J, Attardi G (1981) tRNA punctuation model of RNA processing in human mitochondria. Nature 290:470–474

Page IH (1967) The mosaic theory of arterial hypertension—its interpretation. Perspect Biol Med 10:325–333

Park H, Davidson E, King MP (2003) The pathogenic A3243G mutation in human mitochondrial $tRNA^{Leu(UUR)}$ decreases the efficiency of aminoacylation. Biochemistry 42:958–964. https://doi.org/10.1021/bi026882r

Perez-Cerda C et al (2005) 2-Methyl-3-hydroxybutyryl-CoA dehydrogenase (MHBD) deficiency: an X-linked inborn error of isoleucine metabolism that may mimic a mitochondrial disease. Pediatr Res 58:488–491. https://doi.org/10.1203/01.pdr.0000176916.94328.cd

Powell CA, Nicholls TJ, Minczuk M (2015) Nuclear-encoded factors involved in post-transcriptional processing and modification of mitochondrial tRNAs in human disease. Front Genet 6:79. https://doi.org/10.3389/fgene.2015.00079

Rackham O et al (2016) Hierarchical RNA processing is required for mitochondrial ribosome assembly. Cell Rep 16:1874–1890. https://doi.org/10.1016/j.celrep.2016.07.031

Rauschenberger K et al (2010) A non-enzymatic function of 17beta-hydroxysteroid dehydrogenase type 10 is required for mitochondrial integrity and cell survival. EMBO Mol Med 2:51–62. https://doi.org/10.1002/emmm.200900055

Reid FM, Rovio A, Holt IJ, Jacobs HT (1997) Molecular phenotype of a human lymphoblastoid cell-line homoplasmic for the np 7445 deafness-associated mitochondrial mutation. Hum Mol Genet 6:443–449

Reinhard L, Sridhara S, Hällberg BM (2017) The MRPP1/MRPP2 complex is a tRNA-maturation platform in human mitochondria. Nucleic Acids Res 45(21):12469–12480. https://doi.org/10.1093/nar/gkx902

Richardson A, Berry GT, Garganta C, Abbott MA (2016) Hydroxysteroid 17-beta dehydrogenase type 10 disease in siblings. JIMD Rep 32:25–32. https://doi.org/10.1007/8904_2016_547

Robertson HD, Altman S, Smith JD (1972) Purification and properties of a specific *Escherichia coli* ribonuclease which cleaves a tyrosine transfer ribonucleic acid precursor. J Biol Chem 247:5243–5251

Sanchez MI et al (2011) RNA processing in human mitochondria. Cell Cycle 10:2904–2916. https://doi.org/10.4161/cc.10.17.17060

Schaub MC, Lopez SR, Caputi M (2007) Members of the heterogeneous nuclear ribonucleoprotein H family activate splicing of an HIV-1 splicing substrate by promoting formation of ATP-dependent spliceosomal complexes. J Biol Chem 282:13617–13626

Seaver LH et al (2011) A novel mutation in the HSD17B10 gene of a 10-year-old boy with refractory epilepsy, choreoathetosis and learning disability. PLoS One 6:e27348. https://doi.org/10.1371/journal.pone.0027348

Sen A, Cox RT (2017) Fly models of human diseases: *Drosophila* as a model for understanding human mitochondrial mutations and disease. Curr Top Dev Biol 121:1–27. https://doi.org/10.1016/bs.ctdb.2016.07.001

Sen A, Karasik A, Shanmuganathan A, Mirkovic E, Koutmos M, Cox RT (2016) Loss of the mitochondrial protein-only ribonuclease P complex causes aberrant tRNA processing and lethality in Drosophila. Nucleic Acids Res 44:6409–6422. https://doi.org/10.1093/nar/gkw338

Shafqat N et al (2003) Expanded substrate screenings of human and *Drosophila* type 10 17-β-hydroxysteroid dehydrogenases (HSDs) reveal multiple specificities in bile acid and steroid hormone metabolism: characterization of multifunctional 3α/7α/7β/17β/20β/21-HSD. Biochem J 376:49–60

Shao Z et al (2014) Crystal structure of tRNA m^1G9 methyltransferase Trm10: insight into the catalytic mechanism and recognition of tRNA substrate. Nucleic Acids Res 42:509–525. https://doi.org/10.1093/nar/gkt869

Spradling AC et al (1999) The Berkeley Drosophila genome project gene disruption project: single *P*-element insertions mutating 25% of vital Drosophila genes. Genetics 153:135–177

Sutton VR, O'Brien WE, Clark GD, Kim J, Wanders RJ (2003) 3-Hydroxy-2-methylbutyryl-CoA dehydrogenase deficiency. J Inherit Metab Dis 26:69–71

Taanman J-W (1999) The mitochondrial genome: structure, transcription, translation and replication. Biochim Biophys Acta 1410:103–123. https://doi.org/10.1016/S0005-2728(98)00161-3

Taschner A, Weber C, Buzet A, Hartmann Roland K, Hartig A, Rossmanith W (2012) Nuclear RNase P of *Trypanosoma brucei*: a single protein in place of the multicomponent RNA-protein complex. Cell Rep 2:19–25. https://doi.org/10.1016/j.celrep.2012.05.021

Torroja L, Ortuno-Sahagun D, Ferrus A, Hammerle B, Barbas JA (1998) *scully*, an essential gene of *Drosophila*, is homologous to mammalian mitochondrial type II L-3-hydroxyacyl-CoA dehydrogenase/amyloid-beta peptide-binding protein. J Cell Biol 141:1009–1017

Van Haute L, Pearce SF, Powell CA, D'Souza AR, Nicholls TJ, Minczuk M (2015) Mitochondrial transcript maturation and its disorders. J Inherit Metab Dis 38:655–680. https://doi.org/10.1007/s10545-015-9859-z

Vilardo E, Rossmanith W (2015) Molecular insights into HSD10 disease: impact of SDR5C1 mutations on the human mitochondrial RNase P complex. Nucleic Acids Res 43:5112–5119. https://doi.org/10.1093/nar/gkv408

Vilardo E, Nachbagauer C, Buzet A, Taschner A, Holzmann J, Rossmanith W (2012) A subcomplex of human mitochondrial RNase P is a bifunctional methyltransferase—extensive moonlighting in mitochondrial tRNA biogenesis. Nucleic Acids Res 40:11583–11593. https://doi.org/10.1093/nar/gks910

Wang S et al (2011) Maternally inherited essential hypertension is associated with the novel 4263A>G mutation in the mitochondrial $tRNA^{Ile}$ gene in a large Han Chinese family. Circ Res 108:862–870. https://doi.org/10.1161/CIRCRESAHA.110.231811

Xie X, Dubrovskaya VA, Dubrovsky EB (2011) RNAi knockdown of *dRNaseZ*, the *Drosophila* homolog of *ELAC2*, impairs growth of mitotic and endoreplicating tissues. Insect Biochem Mol Biol 41:167–177. https://doi.org/10.1016/j.ibmb.2010.12.001

Xie X, Dubrovskaya V, Yacoub N, Walska J, Gleason T, Reid K, Dubrovsky EB (2013) Developmental roles of *Drosophila* tRNA processing endonuclease RNase Z^{L} as revealed with a conditional rescue system. Dev Biol 381:324–340. https://doi.org/10.1016/j.ydbio.2013.07.005

Xu F et al (2008) Disruption of a mitochondrial RNA-binding protein gene results in decreased cytochrome *b* expression and a marked reduction in ubiquinol–cytochrome *c* reductase activity in mouse heart mitochondria. Biochem J 416:15

Yang SY et al (2009) Mental retardation linked to mutations in the HSD17B10 gene interfering with neurosteroid and isoleucine metabolism. Proc Natl Acad Sci U S A 106:14820–14824. https://doi.org/10.1073/pnas.0902377106

Yang S-Y, He X-Y, Isaacs C, Dobkin C, Miller D, Philipp M (2014) Roles of 17β-hydroxysteroid dehydrogenase type 10 in neurodegenerative disorders. J Steroid Biochem Mol Biol 143:460–472. https://doi.org/10.1016/j.jsbmb.2014.07.001

Yogev O, Pines O (2011) Dual targeting of mitochondrial proteins: mechanism, regulation and function. Biochim Biophys Acta 1808:1012–1020. https://doi.org/10.1016/j.bbamem.2010.07.004

Zhang Y (2008) I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 9:40. https://doi.org/10.1186/1471-2105-9-40

Zhu HY, Wang SW, Liu L, Li YH, Chen R, Wang L, Holliman CJ (2009) A mitochondrial mutation A4401G is involved in the pathogenesis of left ventricular hypertrophy in Chinese hypertensives. Eur J Hum Genet 17:172–178. https://doi.org/10.1038/ejhg.2008.151

Zschocke J (2012) HSD10 disease: clinical consequences of mutations in the HSD17B10 gene. J Inherit Metab Dis 35:81–89. https://doi.org/10.1007/s10545-011-9415-4

Zschocke J, Ruiter JP, Brand J, Lindner M, Hoffmann GF, Wanders RJ, Mayatepek E (2000) Progressive infantile neurodegeneration caused by 2-methyl-3-hydroxybutyryl-CoA dehydrogenase deficiency: a novel inborn error of branched-chain fatty acid and isoleucine metabolism. Pediatr Res 48:852–855. https://doi.org/10.1203/00006450-200012000-00025

Chapter 4
Intercompartment RNA Trafficking in Mitochondrial Function and Communication

Frédérique Weber-Lotfi and André Dietrich

Abstract Mitochondria currently appear to be a major destination for the RNA trafficking and localization processes that control and coordinate gene expression in living cells. A large set of messenger RNAs derived from the nuclear genome is translated at the mitochondrial surface, while an increasing series of noncoding RNAs has been reported to localize in the organelles, including microRNAs, additional small noncoding RNAs, transfer RNAs, and long noncoding RNAs. These RNA species contribute to mitochondrial functions and control of organellar gene expression, but mitochondria might also store and release noncoding RNAs of nuclear origin having cytosolic targets. Conversely, data have emerged implying that small and long noncoding regulatory RNAs are generated within the organelles from the mitochondrial genome. Some of them nevertheless appeared to localize to the nucleus or were recovered in body fluids. Integrating all reported data leads to an intricate picture of multidirectional RNA trafficking and intercompartment communication that can be related to cellular homeostasis, cell differentiation, pathogenesis, or disease. However, a number of facets in this amazing picture are still a matter of debate, as the mechanisms underlying nucleic acid translocation through the mitochondrial membranes remain difficult to assess and the widespread presence of mitochondrial DNA pseudo-sequences in the nuclear genome can make the origin of some transcripts confusing. A detailed panorama of the reported mitochondrial noncoding RNAs and of the questions raised by the data is developed here in relation to major mitochondrial processes.

Keywords Genetic regulation · Gene expression · Intercompartment cross-talk · Major diseases · Membrane transport · Mitochondria · Noncoding RNA

F. Weber-Lotfi · A. Dietrich (✉)
Université de Strasbourg, CNRS, IBMP UPR 2357, Strasbourg, France
e-mail: andre.dietrich@ibmp-cnrs.unistra.fr

J. Cruz-Reyes, M. Gray (eds.), *RNA Metabolism in Mitochondria*, Nucleic Acids and Molecular Biology 34, https://doi.org/10.1007/978-3-319-78190-7_4

4.1 Introduction

Trafficking and localization of RNAs is a general feature of cellular processes and plays a fundamental role in the efficiency and regulation of gene expression (Buxbaum et al. 2015; Chen 2016; Chin and Lecuyer 2017; Taliaferro et al. 2014). Specific mechanisms involving *cis*-acting sequences and *trans*-acting proteins target coding and noncoding RNAs to the subcellular location where they exert their functions in the frame of an optimized cellular organization. It has progressively appeared that mitochondria are also a destination in these pathways. As is the case for other cytoplasmic compartments, a subset of messenger RNAs (mRNAs) is targeted to the mitochondrial surface for translation (Lesnik et al. 2015; Weis et al. 2013). Such a localization has been shown as well in fungi, mammals, and plants. It is expected to facilitate organellar uptake of the resulting proteins, although it might also help assembly of mitochondrial complexes or improve translation regulation. But mitochondria are a special case, as some RNAs localize to the inside of these organelles, which means that they cross the membranes. While no example of natural import of mRNAs has been documented so far, the extent of noncoding RNAs (ncRNAs) reported to traffic into mitochondria has exploded with the recent accumulation of massive sequencing datasets. The field now extends from microRNAs (miRNAs) to long noncoding RNAs (lncRNAs) through the seminal case of transfer RNAs (tRNAs) (Dong et al. 2017; Kim et al. 2017). Moreover, it is hypothesized that mitochondria store nucleus-encoded ncRNAs that they release to the cytosol upon regulation demands. Not to forget, mitochondria have their own genetic system that produces essential polypeptides, and it has been reported that mitochondrially generated ncRNAs might contribute to the control of organellar gene expression. Further, lncRNAs might be exported from the organelles and exert regulation functions in the nucleus. Altogether, the data imply a complex RNA import-export picture that so far defies the limited knowledge that we have about the possible underlying transport pathways and mechanisms. The challenge is further complicated because the extent of import is highly variable and the low levels of recovery documented for some RNAs in mitochondrial fractions in some cases hinder convincing conclusions. Even the origin of RNAs matching in sequence the mitochondrial DNA (mtDNA) is not necessarily obvious to establish, as numerous pieces of the mitochondrial genome are integrated into the nuclear genome, the so-called nuclear mitochondrial DNA segments (NUMTs) (Richly and Leister 2004; Woischnik and Moraes 2002). The present chapter seeks to give a general panorama of mitochondrial RNA trafficking processes in relation to functions in organellar genetic processes and in regulation pathways. As it will be seen, the above issues make the field rich in controversial data and debate, but also in novel thoughts and prospects.

4.2 Mitochondrial Trafficking of the *MRP* RNA and mtDNA Replication

The mitochondrial genome of mammals is represented as a closed, circular DNA molecule of 16.5 kb. The two strands are depicted as heavy and light, due to diverse buoyant densities resulting from their base composition. The mammalian mtDNA carries only one long noncoding region (D-loop), which contains the transcription promoters for both strands and the origin of replication for the heavy strand (Ori_H). Replication is carried out by DNA polymerase γ. According to the strand-displacement model, mtDNA synthesis on the two strands is asymmetrical and starts first at the Ori_H origin, making the heavy strand the leading strand (Gustafsson et al. 2016). Primase activity is provided by the mitochondrial RNA polymerase (Fusté et al. 2010). RNA production from the light strand promoter (LSP) generates transcripts complementary to the Ori_H region. Processing of these transcripts in turn produces primers to initiate heavy strand replication (Campbell et al. 2012). Based on in vitro approaches, the corresponding processing activity was attributed to a ribonucleoprotein that was called RNase MRP, standing for "RNase for mitochondrial RNA processing" (Chang and Clayton 1987). This enzyme however turned out to be an essential nucleus-encoded endonuclease with various functions in the nuclear and cytosolic compartments (Hernandez-Cid et al. 2012). It processes ribosomal RNAs (rRNAs) and small RNAs in yeast nucleoli and cleaves the B-type cyclin mRNA in temporal asymmetric MRP (TAM) bodies. It is also involved in pre-rRNA processing in plants and protozoa and participates in viral RNA degradation in the cytosol of plant cells.

RNase MRP contains a specific and structured 260–280-nucleotide uncapped RNA encoded by a single-copy nuclear gene that is transcribed by RNA polymerase III. The RNA is considered to act in cleavage catalysis and thus to be indispensable (Topper and Clayton 1990b; Esakova and Krasilnikov 2010). The RNA moiety of RNase MRP is structurally related to that of RNase P (see below Sect. 4.4.3) (Gold et al. 1989; Forster and Altman 1990). In mammalian nuclei, the *MRP* RNA is a source of siRNAs (Maida et al. 2009). As mentioned above, mitochondrial targeting of RNase MRP and of its RNA moiety was put forward very early in mammals and in yeast (Chang and Clayton 1987, 1989; Topper and Clayton 1990a; Bennett and Clayton 1990; Stohl and Clayton 1992) (Fig. 4.1). In both organisms, the enzyme was considered to perform site-specific cleavage of mitochondrial RNA sequences implicated in replication priming (Chang and Clayton 1987; Stohl and Clayton 1992). The *MRP* RNA appeared to be itself processed into a mature form of about 130 nucleotides in mitochondria (Chang and Clayton 1987). However, the significant presence of the *MRP* RNA in mammalian mitochondria was strongly contradicted by a further study in which RNA fractions from highly purified HeLa

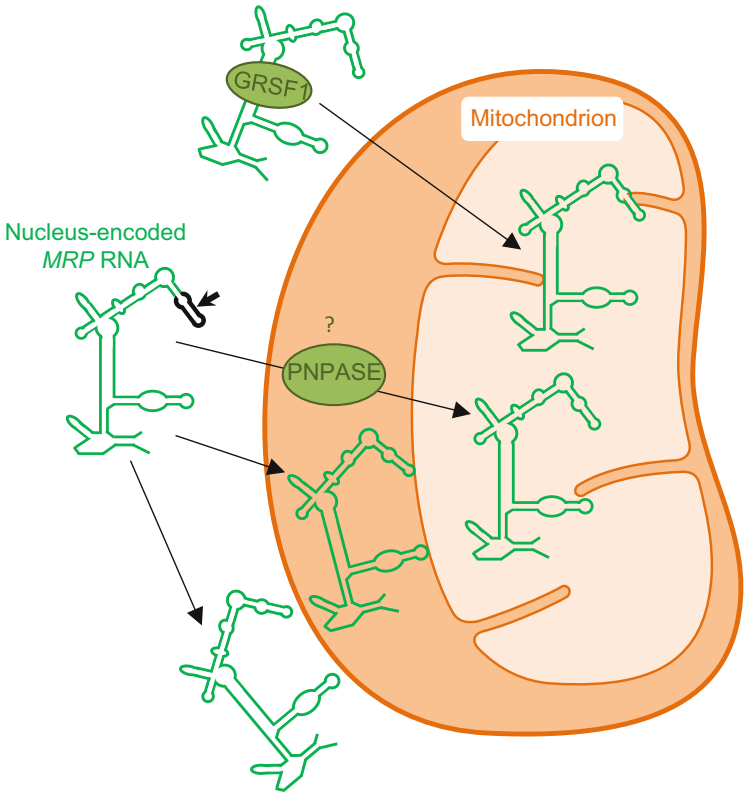

Fig. 4.1 Different options reported for mitochondrial targeting of the nucleus-encoded *MRP* RNA. *MRP* localization to the organelles was proposed to be modulated by the G-rich RNA sequence-binding factor 1 (GRSF1). Import into mitochondria would involve the mitochondrial polynucleotide phosphorylase (PNPASE). It was also put forward that the *MRP* RNA might only reach the mitochondrial intermembrane space or associate with the outer membrane. The stem-loop described as a mitochondrial import determinant (Wang et al. 2010) is highlighted as a thicker black line and pointed by an arrow in one of the *MRP* structures

cell mitoplasts were probed (Kiss and Filipowicz 1992). In these experiments, the level of *MRP* RNA recovered in the final mitoplast fraction was representative of no more than 1 molecule per 100 organelles, i.e., a level that would be unable to sustain RNase MRP activity. A controversial situation thus arose that was further fueled by in situ hybridization experiments pointing to the presence of *MRP* RNA in mammalian mitochondria (Li et al. 1994). In later experiments, additional quantitative analyses implied the presence in mitochondria of 6–15 molecules of *MRP* RNA per human HeLa cell (Puranam and Attardi 2001), which is also quite low considering the hundreds of mitochondria present in a cell (Posakony et al. 1977).

The contradiction was revived again more recently. Extensively purified nucleolar RNase MRP and presumably mitochondrial RNase MRP were characterized in yeast and contained distinct protein components, but an identical RNA moiety

involved in cleavage catalysis (Lu et al. 2010). On the other hand, import of the human RNase MRP RNA moiety into isolated yeast and mammalian mitochondria was reported in a study devoted to a putative role of mitochondrial polynucleotide phosphorylase (PNPASE) in organellar import of RNAs (Wang et al. 2010, 2012a). A predicted 20-nucleotide stem-loop structure in the *MRP* RNA was determined to be a mitochondrial import signal and rendered the non-imported *GAPDH* mRNA able to be taken up into isolated yeast mitochondria (Wang et al. 2010) (Fig. 4.1). Conversely, *MRP* RNA uptake was strongly decreased with liver mitochondria isolated from a mouse with a liver-specific knockout of the PNPASE gene versus mitochondria from wild-type mouse (Wang et al. 2010). The stem-loop structure of the *MRP* RNA was also grafted as an RNA aptamer to mitochondriotropic nanocarriers, so-called MITO-Porters, to optimize the nanovesicle functionality (Yamada et al. 2016). The *MRP* RNA was further detected in mitoplasts in the course of a global analysis of the human mitochondrial transcriptome (Mercer et al. 2011). These reports did not resolve the controversy. The dispute continued on the significance of such observations for the specificity and selectivity of RNA uptake into mitochondria, in the context of a similar situation for the RNase P RNA (see below Sect. 4.4.3) (Rossmanith 2012). In a general search for novel mitochondrially imported RNAs (see Sect. 4.3.3), Cannon et al. (2015) applied to mouse mitochondria a very stringent elimination of cytosolic contamination. As for Kiss et al. with HepG2 cells (Kiss and Filipowicz 1992), they concluded that the *MRP* RNA was not imported into the mitochondrial matrix, only leaving open the possibility that it might associate with the outer membrane or reach the intermembrane space (Fig. 4.1). On the contrary, Noh et al. (2016) reported one more time the presence of the *MRP* RNA in nuclease-treated mitoplasts from WI-38 human fibroblasts. Based on affinity RNA pulldown, mass spectrometry analysis, siRNA silencing, and CRISPR/Cas9-mediated knockdown, they proposed the RNA-binding proteins HuR (human antigen R) and GRSF1 (G-rich RNA sequence-binding factor 1) as modulators of the nuclear export and mitochondrial localization of the *MRP* RNA, respectively (Noh et al. 2016) (Fig. 4.1). The contradictions thus persist. Asking in a more open way the question of the function that the *MRP* RNA would have if present in mitochondria might be a way to make progress.

4.3 ncRNA Trafficking and Regulation of Mitochondrial Gene Expression

Sequencing of whole genomes and transcriptomes has revealed the massive abundance of ncRNAs of different types and sizes that play key roles in all genetic and regulatory processes. Moreover, there is growing evidence that ncRNAs also interact with each other and form complex intertwined networks (Yamamura et al. 2017). After exponential development of the data on ncRNAs in nuclear and cytosolic regulatory mechanisms, the field is now extending to the mitochondrial genetic

compartment, with challenging hypotheses and prospects. Evidence is growing that intercompartment trafficking of ncRNAs takes part in nuclear-mitochondrial communication through retrograde and anterograde signaling (Vendramin et al. 2017) and in the control of most major pathways of mitochondrial metabolism (Baradan et al. 2017; Geiger and Dalgaard 2017). Mitochondrial miRNAs in particular are likely to play important roles in health, disease, and aging (Baradan et al. 2017; Borralho et al. 2015; Rippo et al. 2014; Srinivasan and Das 2015).

4.3.1 Gene Regulation Driven by Imported Nucleus-Encoded miRNAs in Mitochondria

miRNAs are small single-stranded RNAs (ssRNAs) of ~22 nucleotides in size with complex and pleiotropic regulatory roles. Their biogenesis pathway has been well characterized in animals (Breving and Esquela-Kerscher 2010). Transcription of miRNA genes first results in stem-loop pri-miRNA precursors of hundreds to thousands of nucleotides that can carry multiple miRNA sequences. In the nucleus, most pri-miRNAs are recognized and cleaved by the Drosha complex to form pre-miRNAs. The latter are exported from the nucleus to the cytosol by the Exportin-5/Ran-GTP complex. In the cytosol, pre-miRNAs are processed by Dicer and its cofactors, resulting in the release of ~22-nucleotide-long double-stranded RNAs (dsRNAs). These are separated into two ssRNAs, the guide strand and the passenger strand. The guide strand is recovered in the RISC complex that contains one Argonaute protein. In mammals, there are four Argonaute proteins that can bind to the 3′-untranslated region (3′-UTR) of mRNAs. The most abundant Argonaute protein is Argonaute 2 (Ago2). Recruitment of the RISC complex loaded with the guide strand miRNA on its target mRNA blocks translation or triggers degradation (Kim et al. 2016). The mode of action can be selected by proteins of the RISC complex other than Argonaute, such as a member of the GW182 protein family, the RNA-binding protein fragile-X-mental-retardation protein (FMRP), or the decapping activator RCK/p54 (Breving and Esquela-Kerscher 2010). Argonaute proteins and miRNAs have been recovered in the nucleus and in the cytosol. They are associated with P-bodies, stress granules, exosomes, endoplasmic reticulum, Golgi apparatus, lysosomes, and endosomes (Nguyen et al. 2014). Cell-free miRNAs were also found in biological fluids (Makarova et al. 2016).

Global sequence analyses and alignments that run for different mammalian cells pointed out the association of hundreds of indexed nucleus-encoded miRNAs with mitochondria (Bandiera et al. 2011; Barrey et al. 2011; Bian et al. 2010; Das et al. 2012; Dasgupta et al. 2015; Kren et al. 2009; Jagannathan et al. 2015; Mercer et al. 2011; Sripada et al. 2012a; Wang et al. 2015; Wang and Springer 2015). These were named mitomiRs (Bandiera et al. 2011). The first results documenting the presence of nucleus-encoded miRNAs associated with RNase-treated mammalian mitochondria highlighted mitomiRs with no predictable target among mitochondrial RNAs

(Kren et al. 2009; Bian et al. 2010). Indeed, the vast majority of mitomiRs seem to have cytosolic mRNAs as predicted targets and will be discussed below (see Sect. 4.5). However, further analyses highlighted nucleus-originating mitomiRs that can potentially target mtDNA-derived RNAs (Bandiera et al. 2011; Barrey et al. 2011; Dasgupta et al. 2015; Jagannathan et al. 2015; Mercer et al. 2011; Sripada et al. 2012a). Bandiera et al. (2011) identified from HeLa cells a signature of 13 nucleus-encoded miRNAs reproducibly enriched in mitochondrial extracts, out of which 10 had potential mitochondrial targets. Genomic location analyses showed that these miRNAs originated from intragenic or intergenic sequences, most of which were located in regions of mitochondrial pseudogenes in the nuclear genome or in *loci* involved in mitochondrial disorders (Bandiera et al. 2011). In comparison to the usual ~22-nucleotide-long cytosolic miRNAs, they showed unusual sizes varying from 17 to 25 nucleotides, which give them unique thermodynamic features. Also atypical, some mitomiRs potentially target mitochondrial tRNAs or rRNAs (Bandiera et al. 2011; Barrey et al. 2011; Dasgupta et al. 2015; Sripada et al. 2012a). Bioinformatic tools predicted that multiple mitomiRs might target the same mitochondrial mRNA at different positions (Bandiera et al. 2011; Barrey et al. 2011; Dasgupta et al. 2015; Jagannathan et al. 2015) and that a given mitomiR might target several mitochondrial mRNAs, or both mitochondrial and nuclear mRNAs (Bandiera et al. 2011; Dasgupta et al. 2015; Jagannathan et al. 2015; Sripada et al. 2012a).

The functional significance of the presence of miRNAs matching mtDNA sequences in mitochondria has been documented in several cases. Das et al. (2012) reported that nucleus-encoded miR-181c associates with the 3′-UTR of the *COX1* mitochondrial mRNA and Ago2 in organelles (Das et al. 2012) (Fig. 4.2). Overexpression of miR-181c in rat myocyte cultures did not change the *COX1* mRNA level but resulted in a decrease of the COX1 protein in the organelles and a surprising increase of the *COX2* mRNA and COX2 polypeptide. Similarly, in vivo systemic delivery of a miR-181c-encoding plasmid mediated by a nanovector in rats led to a selective mitochondrial complex IV remodeling (Das et al. 2014). Both the *COX1* mRNA and the COX1 protein content were decreased in mitochondria, as well as the *COX2* and *COX3* mRNA levels. Expression of other mitochondrial genes was not significantly affected (Das et al. 2014).

A decreased level of the mitochondrially encoded ATP6 protein was in turn observed upon overexpression of miR-378 in transformed HL-1 cardiomyocyte mouse cells or upon redistribution of miR-378 to the heart interfibrillar mitochondria in diabetic FVB mice (Jagannathan et al. 2015). In mitochondria, miR-378 was found associated with Ago2 and the FRX1 protein, i.e., a mouse structural homolog of FMRP, but the GW182 RISC protein was not detected (Fig. 4.2). Conversely, an enhanced translation of multiple mitochondrial mRNAs was reported in response to miR-1, which is specifically induced during myogenesis (Zhang et al. 2014). The levels of the ND1 and COX1 mtDNA-encoded proteins were increased by more than 15-fold during differentiation of C2C12 mouse cells from myoblasts to myotubes, while the levels of the corresponding mRNAs were invariant. Again, Ago2, but not GW182, was detected in mitochondria (Fig. 4.2). Finally, a specific increase of the

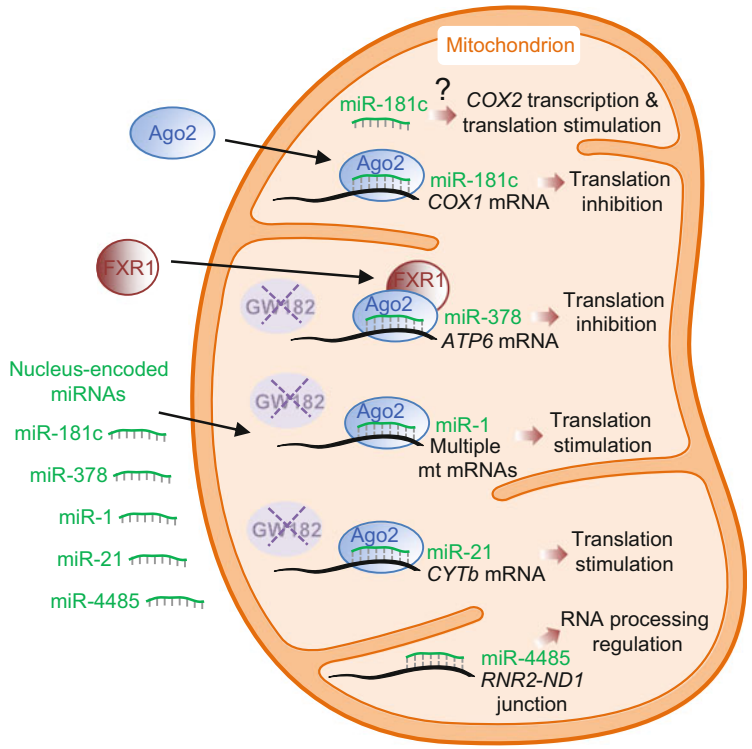

Fig. 4.2 Reported effects of individual nucleus-encoded miRNAs on mitochondrial genetic processes. miRNAs are represented in association with their target. Whether the effect of miR-181c on *COX2* is primary or secondary remains to be established. Ago2 was found in mitochondria in most cases. In the case of miR-378, miR-1, and miR-21, GW182 was searched for but was not detected in mitochondrial fractions. Ago2 is the Argonaute 2 protein; GW182 is a protein containing multiple GW repeats that is essential for miRNA-mediated gene silencing; FXR1 is the fragile-X-mental-retardation-related protein 1; *COX1*, *COX2*, *ATP6*, *CYTb*, and *ND1* are the mRNAs coding for subunit 1 of cytochrome oxidase, subunit 2 of cytochrome oxidase, subunit 6 of ATP synthase, and cytochrome *b* and subunit 1 of NADH Coenzyme Q oxidoreductase, respectively; *RNR2* is the mitochondrial 16S rRNA

CYTb protein, without change at the mRNA level, was observed in H9C2 rat cardiomyoblast cells and in human HK2, HEK293, or HUVEC cells transfected with a miR-21 mimic (Li et al. 2016). Also, Ago2 showed an increased association with the *CYTb* mitochondrial mRNA in miR-21-transfected cells (Fig. 4.2). Notably, GW182 knockdown prevented translational repression of the miR-21 cytosolic target PTEN, but miR-21-triggered enhancement of *CYTb* mRNA translation was maintained (Li et al. 2016). Sripada et al. (2017) localized miR-4485 in mitochondria from human cells. A dynamic association was highlighted, with various levels observed under different stress conditions. miR-4485 was found to bind to mitochondrial 16S rRNA, to regulate the processing of pre-rRNA at the 16S rRNA-*ND1* junction, and to affect translation of the downstream transcripts. Cell transfection

with a miR-4485 mimic affected mitochondrial Complex I activity, ATP production, ROS levels, caspase-3/7 activation, and apoptosis (Sripada et al. 2017).

There is no information so far on putative nucleus-encoded miRNAs that would regulate gene expression in plant mitochondria. Using a computational approach, Kamarajan et al. (2012) predicted seven potential mitochondrial miRNA targets in plants, but no supporting experimental evidence has been presented.

The above series of functional results implies that some nucleus-encoded miRNAs can indeed be translocated into mitochondria and can gain access to mRNAs inside the organelles. Their impact seems to be mostly, but not exclusively, on mitochondrial mRNA translation. However, they appear to either enhance or inhibit translation, depending on the target and the context. Also, while several groups documented the presence of Ago2 in mammalian mitochondria, the absence of important cofactors of the cytosolic RNA silencing pathway, like the GW182 proteins (Pfaff and Meister 2013), was repeatedly reported (Jagannathan et al. 2015; Li et al. 2016; Zhang et al. 2014). The molecular mode of action of miRNAs in mitochondria is thus likely to differ significantly from the known cytosolic mechanisms. Nevertheless, the abovementioned miRNAs appear to be part of important mitochondrial regulation processes. Regulation of COX1 by miR-181c alters mitochondrial function, leads to propensity for heart failure (Das et al. 2014), and connects with the nuclear factor (erythroid-derived 2)-like 2 (NFE2L2/NRF2) pathway in cancer (Jung et al. 2017). Following diabetic insult, miR-378 is redistributed to the interfibrillar mitochondria in the heart, where it participates in alteration of ATP synthase functions through downregulation of ATP6 expression (Jagannathan et al. 2015). This pathway is part of a more general scheme involving the organelles, as miR-378 is also involved in the control of mitochondrial fatty acid oxidation via the carnitine system. Aberrant expression of miR-378 in cancer cells contributes to the deregulation of the carnitine cycle (Valentino et al. 2017). The mitochondrial enzyme carnitine *O*-acetyltransferase in particular is a target controlled by miR-378. KO mice for miR-378 are resistant to high-fat diet-induced obesity and show enhanced mitochondrial fatty acid metabolism (Carrer et al. 2012). Various aspects of the implication of miRNAs, including mRNA targeting inside the organelles, in the regulation of mitochondrial energy metabolism have been discussed, with a special interest for heart failure processes (Pinti et al. 2017). The action of miR-1 on mitochondrial translation contributes to the changes in bioenergetics mechanisms that are part of the myogenesis cell differentiation program. Notably, while stimulating translation of multiple mRNAs in mitochondria, miR-1 is able to repress its nucleus-encoded cytosolic mRNA targets (Zhang et al. 2014). The positive function of miR-21 in mitochondrial translation is involved in the control of blood pressure in mammals (Li et al. 2016). Finally, mitochondrially targeted miR-4485 negatively affects the tumorigenic potential of breast cancer cells in cell culture and in a mouse xenograft model. It might potentially act as a tumor suppressor by downregulating mitochondrial RNA processing and mitochondrial functions (Sripada et al. 2017).

The presence of miRNAs in mitochondria implies that they can cross mitochondrial membranes. However, it has been noted that the level of some mitomiRs is depleted in mitoplasts, i.e., organelles recovered after rupture of the outer membrane,

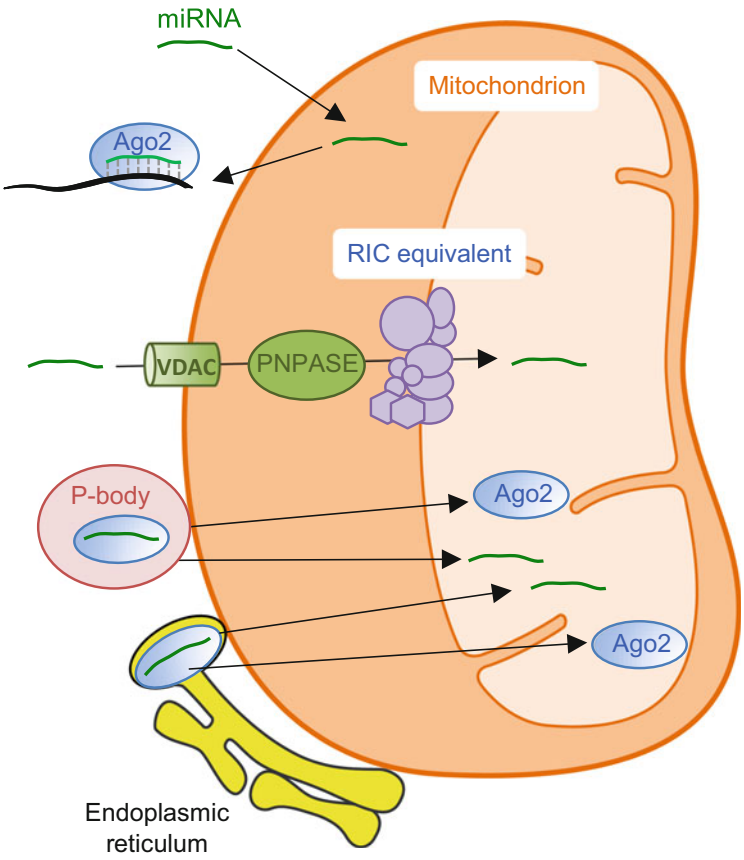

Fig. 4.3 Different options proposed for Ago2-independent mitochondrial trafficking of nucleus-encoded miRNAs. Some miRNAs would enter the intermembrane space for storage and would be released into the cytosol on demand to regulate their target. Mediators suggested for nucleus-encoded miRNA import include the VDAC in the outer membrane, the PNPASE in the intermembrane space and a complex equivalent to the RIC complex in the inner membrane. Interaction of P-bodies or endoplasmic reticulum with mitochondria could also facilitate the independent import of miRNAs and Ago2 into mitochondria. Ago2 is the Argonaute 2 protein; VDAC is the voltage-dependent anion channel; PNPASE is the mitochondrial polynucleotide phosphorylase; RIC is the mitochondrial RNA import complex reported in *L. tropica*

in comparison to intact mitochondria (Mercer et al. 2011; Sripada et al. 2012a). It can thus be assumed that some miRNAs can cross only the outer membrane and remain in the intermembrane space, which would be an easier way for those that are supposed to return to the cytosol to meet their target (see Sect. 4.5) (Fig. 4.3). Conversely, miRNAs efficiently recovered in mitoplasts should be considered to

translocate through both mitochondrial membranes. Several hypotheses have been put forward to account for the uptake mechanism of these miRNAs into mitochondria. Bandiera et al. (2011) proposed that the structural features of mitomiRs could promote their entry into mitochondria (Fig. 4.3). The use of previously identified RNA transport mediators has also been suggested, such as the voltage-dependent anion channel (VDAC) localized in the outer membrane and known to be involved in mitochondrial import of tRNAs in plant cells (see Sect. 4.4.1.2) (Bandiera et al. 2013; Salinas et al. 2006, 2014) (Fig. 4.3). Another factor could be the PNPASE located in the mitochondrial intermembrane space (Bandiera et al. 2013; Geiger and Dalgaard 2017; Sripada et al. 2012b), which was reported to play a role in regulating the import of nucleus-encoded RNAs into mitochondria (Wang et al. 2010) (see Sects. 4.2 and 4.4.3) (Fig. 4.3). Sripada et al. (2012b) also thought about a specific mechanism similar to the RIC complex described as located in the mitochondrial inner membrane of *Leishmania tropica* and supposed to mediate tRNA import into mitochondria in Kinetoplastida (Mukherjee et al. 2007) (see Sect. 4.4.1.1) (Fig. 4.3). Finally, Makarova et al. (2016) proposed that interactions with P-bodies and the endoplasmic reticulum could facilitate the import of miRNAs and Argonaute proteins into mitochondria (Fig. 4.3).

As Ago2 is dual-localized in the cytosol and in mitochondria, Ago2-dependent mechanisms have been suggested for organellar import of miRNAs (Bandiera et al. 2013; Geiger and Dalgaard 2017; Srinivasan and Das 2015; Sripada et al. 2012b; Zhang et al. 2014). However, the pathway involved in mitochondrial import of Ago2 itself is not clear. Fusion with the predicted mitochondrial targeting sequence (MTS) of the Ago2 protein (Bandiera et al. 2011) failed to direct the green fluorescent protein (GFP) reporter to mitochondria (Zhang et al. 2014). Communication between mitochondria and the endoplasmic reticulum was proposed as an alternative to the regular MTS-driven pathway for Ago2 import (Zhang et al. 2014) (Fig. 4.4). Indeed, the endoplasmic reticulum was reported to be a central nucleation site for the assembly of RNA silencing complexes (Barman and Bhattacharyya 2015; Stalder et al. 2013). Conversely, Srinivasan and Das (2015) still suggested that Ago2 could have a pivotal role in transporting ncRNAs into mitochondria via the regular TOM and TIM protein import channels (Fig. 4.4). The authors further speculated that Ago2 might also export to the cytosol miRNAs stored in mitochondria (see Sect. 4.5) (Fig. 4.4). In a recent study, Ago2 and PNPASE were co-immunoprecipitated from mitochondrial pellets, and the overexpression of PNPASE led to an increase in the miR-378 level in the mitochondria of HL-1 cardiomyocyte mouse cells (Shepherd et al. 2017). Strikingly, knockdown of the protein did not significantly alter the mitochondrial miR-378 level. While the results show an association of Ago2 and PNPASE, establishing whether the complex is inside mitochondria will require further experiments using protease treatment of mitochondria or preparation of mitoplasts (Fig. 4.4).

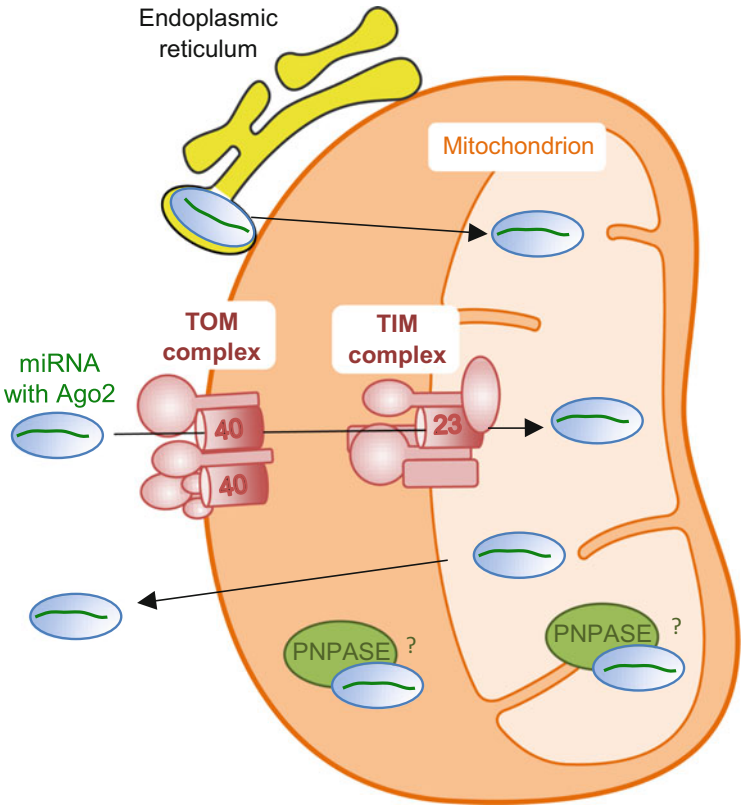

Fig. 4.4 Different options proposed for Ago2-dependent mitochondrial trafficking of nucleus-encoded miRNAs. The endoplasmic reticulum was proposed to be involved in mitochondrial import of miRNAs in complex with Ago2. Alternatively, Ago2 could use the regular TOM and TIM complexes to transport miRNAs into mitochondria. It was also suggested that Ago2 might export stored miRNAs from mitochondria to the cytosol. Ago2 and PNPASE were co-immunoprecipitated from mitochondrial fractions. Ago2 is the Argonaute 2 protein; PNPASE is the mitochondrial polynucleotide phosphorylase; TOM and TIM are the regular protein translocases of the outer and inner membranes

4.3.2 Gene Regulation Driven by mtDNA-Encoded miRNAs in Mitochondria

As discussed above, mitomiR sequences mapping to the mtDNA could mostly be attributed to indexed nucleus-encoded miRNAs, but whether some derive from mtDNA-encoded transcripts is an obvious question (Barrey et al. 2011; Sripada et al. 2012a). Giving a clear-cut answer is complex, due to the presence of numerous pieces of mitochondrial sequences, i.e., NUMTs, in the mammalian nuclear genome, but the potential exists. Both strands of the mammalian mtDNA are entirely transcribed, generating large polycistronic RNAs. H-strand transcription is initiated from

two differentially regulated sites in the D-loop, HSP1 and HSP2 (Asin-Cayuela and Gustafsson 2007). The HSP1 promoter produces a transcript that covers only the 2 rRNA genes, but HSP2 yields fully coding transcripts that cover almost the entire H-strand, including the 2 rRNA genes, 15 tRNA genes, and 12 protein genes. L-strand transcription is initiated from the LSP promoter and produces a transcript that carries 7 tRNAs and the *ND6* mRNA. In this context, processing of the primary transcripts can potentially yield antisense RNAs to any of the mitochondrial rRNAs, tRNAs, or mRNAs. In particular, processing of the transcripts initiated from the LSP promoter releases a set of large noncoding sequences that are antisense to most of the mitochondrial genes. Analysis of the human mitochondrial transcriptome annotated a series of novel small ncRNAs potentially expressed from distinct mtDNA loci, but the majority seemed to derive from mitochondrial tRNA cleavage (Mercer et al. 2011). Numerous small noncoding transcripts (12–137 nucleotides) mapping to the mitochondrial genome were identified in a further deep sequencing study of cytosolic and mitochondrial RNA libraries from mouse and human (Ro et al. 2013). Comparison with the transcriptome of mtDNA-depleted ρ0 cells supported the conclusion that these potential ncRNAs were encoded by the mitochondrial genome. Analyses of their 5′-ends suggested that they were products specifically generated from larger transcripts by ribonucleases rather than residual turnover by-products. Notably, the patterns were different among organs and between species. When some of the candidates were transfected into NIH 3 T3 mouse cells, an increased expression of their putative mitochondrial target RNAs was observed (Ro et al. 2013). However, the possibility was not considered that in such cell transfection assays, small RNAs presumed to be normally generated inside mitochondria have in this case to be targeted from the cytosol into the organelles to reach their targets. Also puzzling, production of the presumed mtDNA-encoded small RNAs was proposed to involve the Dicer RNA interference ribonuclease, but Dicer was not detected in the mitochondrial fractions from human HEK293T cells. Ago2 was also not found in the organellar fractions in that study (Ro et al. 2013).

Barrey et al. (2011) identified 25 potential pre-miRNAs in mitochondria, which in turn agrees with the idea of a mitochondrial miRNA biosynthesis pathway. In addition, Shinde and Bhadra (2015) used in silico analysis of the mitochondrial human genome to predict pre-miRNA-like hairpin structures and find miRNA target sites. Six suitable pre-miRNA and mature miRNA candidates were selected. The size of the predicted miRNAs varied between 19 and 22 nucleotides. These pre-miRNAs and miRNAs were detected by quantitative northern analysis and RT-qPCR in enriched mitochondrial fractions containing low cytosolic mRNA contamination. For preliminary functional analyses, the wild-type or a mutated 3′-region of the mitochondrial 16S rRNA gene was cloned downstream of the luciferase gene. The fusion construct was expressed in human HEK-293 cells co-transfected with candidate mitochondrial miRNAs. Downregulation of luciferase was observed when using the wild-type fusion, whereas fusing the mutated 3′-region had no consequence. The assays suggested that the mitochondrial sequence tested was a potential target for two out of the six miRNA candidates analyzed. Further sequence and function analyses should strengthen these observations.

The large size (commonly 200–700 kb) and low gene density of plant mitochondrial genomes (Gualberto et al. 2014) theoretically provide a wide potential for producing ncRNAs, while the complexity of transcriptional and posttranscriptional processes (Bonen 2008; Brown et al. 2014; Holec et al. 2006, 2008; Sun et al. 2016; Takenaka et al. 2013) indicates multiple regulation mechanisms. Nevertheless, the existence of mtDNA-derived ncRNAs with specific functions in plant mitochondria remains to be assessed (Rurek 2016). Of particular interest is the case of the angiosperm genus *Silene*. Some species in this genus have the largest mitochondrial genomes ever characterized and sequenced, reaching over 11 Mb in size (Sloan et al. 2012; Wu et al. 2015a). Resulting from massive gain and duplication of noncoding sequences, these enormous mtDNAs are fragmented into dozens of circular-mapping mitochondrial chromosomes, some of which contain no identifiable genes but still are maintained. Whether the "empty" chromosomes contain novel and unannotated functional elements, in particular genes for functional ncRNAs, was investigated through RNA-seq of enriched mitochondrial transcripts of *Silene noctiflora* transcripts (Wu et al. 2015b). The overall frequency of 17- to 25-nucleotide RNA reads mapping to the mitochondrial genome was very low. No candidates for small regulatory RNAs antisense to annotated mitochondrial genes were identified. A large number of localized areas in unannotated regions, including "empty" chromosomes, showed high transcript abundance; however, only four candidates for functional small ncRNAs mapping outside of regions containing annotated genes were detected, and none of them exhibited significant similarity to characterized small RNAs in miRBase (Wu et al. 2015b). Deep sequencing profiling of *Arabidopsis thaliana* small RNAs in the 18- to 30-nucleotide range under high- or low-temperature stress detected differential sets of unique reads mapping to the mitochondrial genome and accounting for 0.40 (normal temperature), 0.37 (low temperature), and 0.43 (high temperature) percent of the total unique reads, respectively (Baev et al. 2014). Whether mitochondrially encoded small RNAs indeed contribute to temperature stress response in plants remains to be considered in more detail.

4.3.3 Gene Regulation Driven by Nucleus-Encoded lncRNAs in Mitochondria

lncRNAs are transcripts of 200 nucleotides and more in length. A vast range of lncRNAs is involved in all aspects of eukaryotic genetic and regulation processes (Bunch 2017; Long et al. 2017; Wang et al. 2017). Recent data suggest that some nucleus-originating lncRNAs might be targeted to mitochondria. In a general search for novel mitochondrially imported RNAs, Cannon et al. (2015) subjected mitochondria isolated from mouse liver to a stepwise removal of the outer membrane with digitonin. RNase digestion was carried out at each step. Digitonin treatment was titrated in increasing quantities to ensure full elimination of the mitochondrial

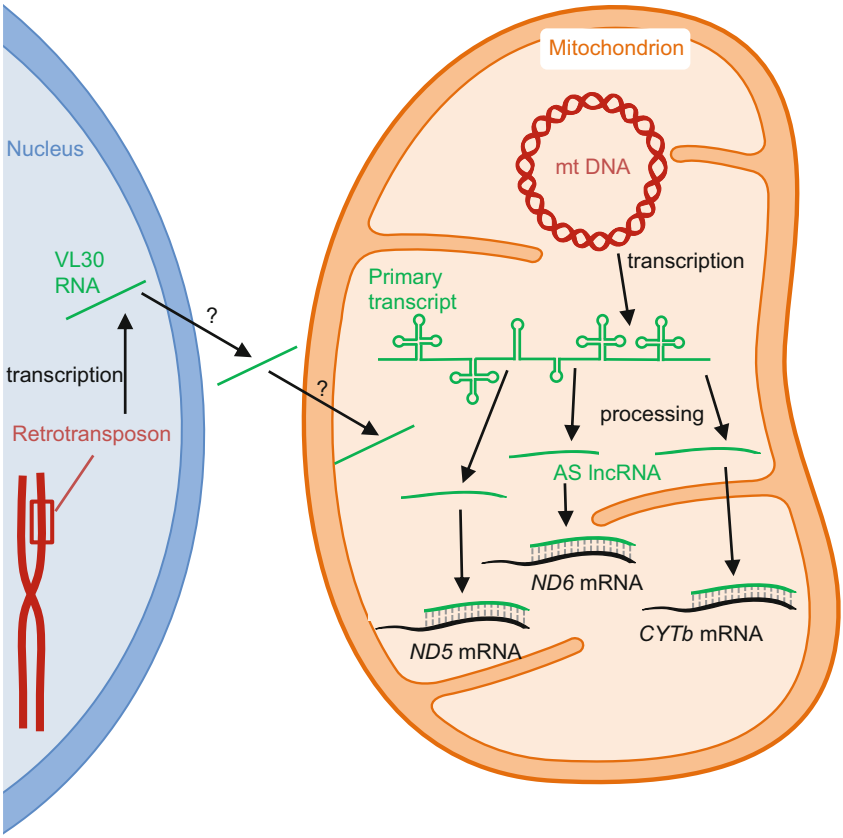

Fig. 4.5 Origin of lncRNAs found in mammalian mitochondria. Nucleus-encoded transcripts of the VL30 retroelement were recovered in stringently generated mitoplasts. Processing of mitochondrial primary transcripts was reported to generate antisense (AS) lncRNAs that can form duplexes with the corresponding sense mitochondrial mRNAs. *ND5*, *ND6*, and *CYTb* are the mRNAs coding for subunits 5 and 6 of the NADH-ubiquinone oxidoreductase and cytochrome *b*, respectively

outer membrane at the final step. In such stringent conditions, only RNAs deriving from VL30 retroelements were identified in the final mitochondrial matrix fraction (Fig. 4.5). The RT-qPCR assays used did not give access to RNA sizes, but VL30 RNAs are known to be in the range of 1 to 11 kb (French and Norton 1997). What function(s) such large RNAs might have in mitochondria and how they can translocate into the organelles remain to be uncovered. Notably, it was reported that VL30 retrotransposition mediates cell death via mitochondrial and lysosomal damage, pointing to a role of retrotransposition as a nuclear signal activating a mitochondrial-lysosomal cross-talk in triggering cell death (Noutsopoulos et al. 2010). The results of these studies also implied that the classical approach of treating

an unpurified mitoplast pellet with RNase might be insufficient to completely eliminate RNAs that are not inside mitochondria (Cannon et al. 2015).

The recently annotated SAMMSON lncRNA is involved in melanogenesis. It promotes mitochondrial targeting of P32, a regulator of mitochondrial homeostasis and metabolism. Silencing SAMMSON disrupted mitochondrial functions in a cancer cell-specific manner (Leucci et al. 2016). Interestingly, RNA-FISH experiments showed that the SAMMSON lncRNA co-localized to a large extent with mitochondria in melanoma cell lines (Leucci et al. 2016). Finally, *SRA1* is a bifunctional gene, both producing a lncRNA and encoding a conserved protein (Leygue 2007). The *SRA* lncRNA is a target for a number of RNA-binding proteins, including the *SRA* stem-loop-interacting RNA-binding protein (SLIRP). Based on the reported observation that SLIRP is predominantly a mitochondrial protein (Colley and Leedman 2009), it was speculated that the *SRA* lncRNA might have a role in both the nucleus and the mitochondria (Dong et al. 2017).

4.3.4 Gene Regulation Driven by mtDNA-Encoded lncRNAs in Mitochondria

As mentioned above, the mammalian mitochondrial genome is extremely compact, but expression from the HSP and LSP promoters generates full transcripts from both strands and opens the possibility to release ncRNAs. Besides putative miRNAs, identification of candidate mtDNA-encoded lncRNAs was reported. Strand-specific deep sequencing analyses run on the transcriptome of RNase-treated HeLa cell mitochondria highlighted three putative lncRNAs complementary to the *ND5*, *ND6*, and *CYTb* mitochondrial mRNAs (Rackham et al. 2011) (Fig. 4.5). It was estimated by RT-qPCR that these lncRNAs were 58%, 34%, and 14% as abundant as their complementary coding *ND5*, *ND6*, and *CYTb* mRNAs, respectively, and that they contained no significant open reading frame. Structural analyses showed that they all formed intermolecular duplexes, suggesting that they may have a functional role. Expression of these lncRNAs appeared to be regulated by nucleus-encoded mitochondrial proteins involved in RNA processing, a further support for their significance. Knockdown of mitochondrial RNase P proteins 1 and 3 (MRPP1 and MRPP3) caused a dramatic decrease in the level of the three lncRNAs, with a lower effect for MRPP3. A subtler negative effect was observed upon knockdown of the mitochondrial RNase Z ELAC2, whereas impairing the RNA-binding protein PTCD1 was associated with a specific decrease of the lnc*ND5* RNA. Knockdown of PTCD2 resulted in a decrease of lnc*ND5* and lnc*ND6* (Rackham et al. 2011). Notably, PTCD1 and PTCD2 are pentatricopeptide repeat (PPR) proteins (Lightowlers and Chrzanowska-Lightowlers 2008). Finally, the fact that the abundance of these lncRNAs varied in different cell lines and tissues further supported the idea that they contribute to the regulation of mitochondrial gene expression.

Additional putative mitochondrial lncRNAs were identified upon deep sequencing of left ventricle myocardial transcripts from patients with severe heart failure (Yang et al. 2014). Strikingly, reads mapping to the mitochondrial genome were highly abundant in the total lncRNA population, with nine putative mitochondrial lncRNAs accounting for 71% of the total cardiac lncRNA read counts. Analysis of the abundance of these potentially mtDNA-encoded lncRNAs across different disease states, together with that of regulatory factors, revealed a significant negative correlation with the nucleus-encoded transcription factors GABPA and NRF1, as well as with the coactivator of mitochondrial biogenesis PPRC1. It was proposed that a decreased mitochondrial transcript abundance in failing myocardium might trigger an upregulation of nucleus-encoded regulators of mitochondrial biogenesis (Yang et al. 2014). Again, the regulatory processes described seem to support the significance of these mitochondrial lncRNA candidates, but whether they have themselves a role in the regulation of mitochondrial gene expression remains to be investigated.

An early cDNA library survey in *A. thaliana* identified 5 putative mitochondrial ncRNAs with sizes of 162 to 300 nucleotides (Marker et al. 2002). Four of these ncRNAs mapped to intergenic regions, surrounded by open reading frames of undetermined significance but also by tRNA or respiratory complex subunit genes. The fifth candidate spanned the *nad1* exon 5 3′-end and its downstream region, with a tissue-specific expression mostly occurring in roots (Marker et al. 2002). Four of these transcripts identified from a total cDNA library also mapped to the *A. thaliana* nuclear genome, which contains massive insertions of mitochondrial sequences, so that their origin remained uncertain. A 500-nucleotide mitochondrial transcript actively expressed from an intergenic region was also highlighted during analysis of mtDNA transcription and organelle RNA stability in *A. thaliana* (Holec et al. 2006). Notably, this transcript turned out to be edited, which suggests an authentic mitochondrial origin, although the possibility cannot be discounted that RNAs imported into mitochondria might be edited if the appropriate sequence context happens to be present. As mentioned above (Sect. 4.3.2), RNA-seq analysis of the transcripts deriving from the enormous 7 Mb mitochondrial genome of *S. noctiflora* identified a large number of localized areas in intergenic regions, including "empty" chromosomes, showing high transcription levels (Wu et al. 2015b). Whether these are a source of functional long or intermediate size ncRNAs remains to be investigated, but the idea is supported by the fact that many of these transcribed intergenic sequences carry C to U editing sites.

Other extended plant mitochondrial transcriptomes have been generated through deep sequencing, but the possible existence of functional ncRNAs in the organelles was not considered (Fang et al. 2012; Grewe et al. 2014; Grimes et al. 2014; Islam et al. 2013; Picardi et al. 2010; Shearman et al. 2014). However, again a number of editing sites were detected in noncoding and unannotated regions. In particular, RNA-seq analysis of the tobacco (*Nicotiana tabacum*) mitochondrial transcriptome detected 73 editing sites in intergenic regions (Grimes et al. 2014). Finally, a total of 68 ncRNA candidates presumably derived from the mitochondrial or chloroplastic genome was reported upon global sequencing of the 50- to 500-nucleotide RNAs of

rice (*Oryza sativa*) (Liu et al. 2013). Sequencing of *A. thaliana* RNAs in the 50–300-nucleotide range in turn highlighted 49 noncoding transcripts mapping to the mitochondrial genome (Wang et al. 2014).

4.4 RNA Trafficking and Mitochondrial Translation

4.4.1 Mitochondrial Trafficking of Nucleus-Encoded tRNAs to Support Translation

The set of mtDNA-encoded tRNAs is often insufficient for translation of all codons. Some of the mitochondrial tRNAs are encoded by the nuclear genome and imported from the cytosol into the organelles, where they function in protein synthesis. The concept of tRNA import into mitochondria appeared 50 years ago with the analysis of the organellar tRNA population in the ciliated protozoan *Tetrahymena pyriformis* (Suyama 1967). Far from the initial skepticism, it is now accepted as an essential process of mitochondrial biogenesis and is widespread in all eukaryotes. The number and identity of the mitochondrially imported isoacceptors, which generally partition between the cytosol and the organelles, depend on the species. The targeting and translocation mechanisms remain difficult to assess but appear to differ among the organisms (Salinas et al. 2008; Schneider 2011; Rubio and Hopper 2011; Sieber et al. 2011; Wang et al. 2012a; Salinas-Giegé et al. 2015).

4.4.1.1 Mitochondrial Trafficking of Nucleus-Encoded tRNAs in Protists

Kinetoplastida (*Trypanosoma*, *Leishmania*) and Apicomplexa (*Plasmodium*, *Toxoplasma*) represent an extreme case, as their mtDNA carries no tRNA gene and thus all mitochondrial tRNAs are imported from the cytosol (Hancock and Hajduk 1990; Kapushoc et al. 2002; Tan et al. 2002a; Esseiva et al. 2004; Sharma and Sharma 2015). However, this is not necessarily the case in other protists. The *T. pyriformis* mtDNA carries 8 tRNA genes encoding 7 distinct tRNAs (Chiu et al. 1975; Suyama 1986; Burger et al. 2000), while the bacteria-like and gene-rich mitochondrial genomes of core jakobids (Excavata) contain up to 29 tRNA genes, with 27 for *Reclinomonas americana* (Burger et al. 2013). Notably, even in extreme cases where the mtDNA carries no tRNA gene, the import process is selective, i.e., not all cytosolic tRNAs are recovered in mitochondria. The reasons for that can be multiple (Sieber et al. 2011). Some tRNAs are not needed, due to the prokaryotic features of mitochondrial translation. Others would be inefficient or deleterious, due to organellar deviations from the universal genetic code. On the other hand, a tRNA can also be imported and subsequently adapted to the mitochondrial genetic code through editing of the anticodon (Charriere et al. 2006). The initiator tRNA is a

special example. Mitochondrial translation requires a formylated initiator $tRNA^{Met}$. The cytosolic initiator $tRNA^{Met}$ cannot be formylated, and accordingly it is not present in mitochondria. Instead, a fraction of the imported cytosolic elongator $tRNA^{Met}$ in *Trypanosoma brucei* can be formylated in the single mitochondrion, so that this tRNA can be used for both elongation and initiation of translation in mitochondria (Tan et al. 2002b). The extent of mitochondrial localization of the tRNAs shared with the cytosol varies in *Leishmania* and *Trypanosoma*, but there seems to be no correlation with the cytosolic abundance, with the organellar codon usage, or with the life cycle (Shi et al. 1994; Kapushoc et al. 2002; Tan et al. 2002a; Cristodero et al. 2010). Attempting to take advantage of the tRNA import pathway, Sbicego et al. (1998) exchanged the natural intron in $tRNA^{Tyr}$ for synthetic sequences and transformed splicing-deficient *Leishmania tarentolae* cell lines with the corresponding gene constructs. Modified tRNAs containing up to 38 nucleotides of additional sequence were transported into the mitochondria.

Structural determinants for tRNA import have been analyzed through different in vitro and in vivo approaches. For *Leishmania*, the reported features included the D-arm, the anticodon, the T-arm, and the variable loop, depending on the tRNA considered (Bhattacharyya et al. 2002). The D-arm was especially highlighted for $tRNA^{Tyr}$ (Mahapatra et al. 1998) and was taken as an import signal in further work (see below). In *Tetrahymena thermophila*, the UUG anticodon of $tRNA^{Gln}$ was shown to be necessary and sufficient for mitochondrial import (Rusconi and Cech 1996). The nature of the T-stem 51:63 base pair determined mitochondrial localization in *T. brucei*, while the U51:A63 base pair of the initiator $tRNA^{Met}$ constituted an import antideterminant for this tRNA (Crausaz Esseiva et al. 2004). In line with such observations, the elongation factor eEF1α was shown to be essential for the targeting of tRNAs to the mitochondrial surface in *T. brucei* (Fig. 4.6) and the U51:A63 base pair happened to be the major antideterminant preventing binding of the non-imported initiator $tRNA^{Met}$ to eEF1α (Bouzaidi-Tiali et al. 2007). Membrane translocation of the cytosolic tRNAs targeted to the mitochondrial surface in protozoa seems to occur through a mechanism distinct from protein import (Schneider 2011; Rubio and Hopper 2011). Characterization of an outer membrane receptor (called TAB) for tRNA import into *Leishmania tropica* mitochondria was reported in early studies (Mahapatra and Adhya 1996) (Fig. 4.6), but this observation has had no further follow-up so far. As to tRNA translocation through the inner membrane, a large RNA import complex (RIC), composed of three mtDNA-encoded subunits and eight nucleus-encoded subunits, was isolated from *L. tropica* mitochondria and further analyzed (Bhattacharyya et al. 2003; Chatterjee et al. 2006; Mukherjee et al. 2007; Koley and Adhya 2013) (Fig. 4.6). The RIC complex allowed the import of the human $tRNA^{Lys}$ into isolated human mitochondria (Mahata et al. 2005). It was also reported that isolated RIC complex could enter human cells by a caveolin-1-dependent pathway and subsequently support mitochondrial import of endogenous tRNAs (Mahata et al. 2006; Mukherjee et al. 2014). However, the data on the RIC became the subject of another open controversy in the field. Aspects of the work raised an editorial expression of concern (Schekman 2010). Also, no other group could find evidence for the existence of such a complex in *Leishmania* or

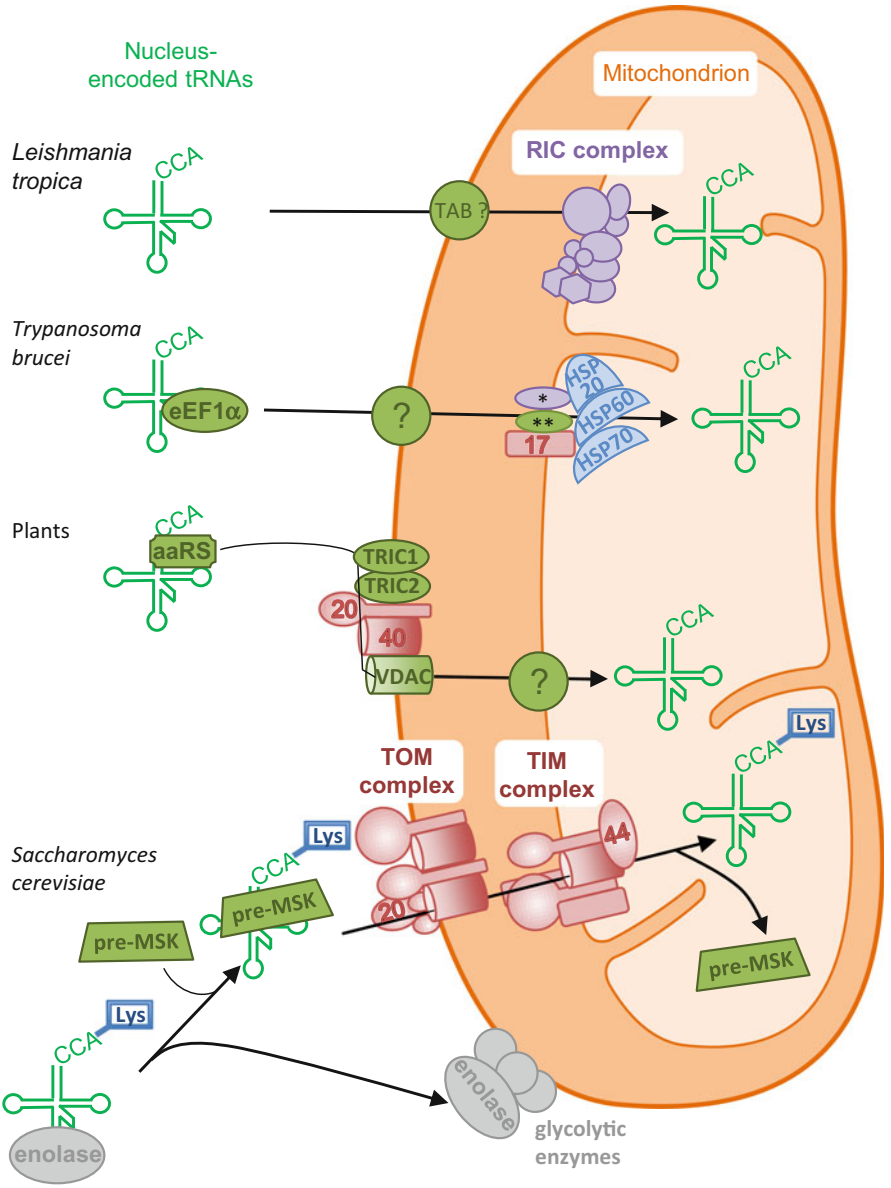

Fig. 4.6 Models of mitochondrial tRNA import mechanisms in different organisms. Proteins of the protist RIC complex are represented in purple. The proteins of the yeast mitochondrial protein import machinery, the translocases of the outer and inner membranes (TOM and TIM), are in brown. Heat shock proteins are in blue. TAB is an RNA-binding protein called tubulin antisense-binding protein; eEF1α is the eukaryotic translation elongation factor 1 alpha; single and double asterisks are the *Trypanosoma* tRNA-binding membrane proteins Tb09.v1.0420 and Tb11.01.4590, respectively; Tb09.v1.0420 is a homolog of subunit 11 of the RIC complex; aaRS stands for aminoacyl-tRNA synthetase; TRIC1 and TRIC2 are tRNA import components identified in *A. thaliana*; VDAC is the voltage-dependent anion channel; pre-MSK1 is the precursor of the mitochondrial lysyl-tRNA synthetase; enolase is a phosphopyruvate hydratase

Trypanosoma. Participation of some of the described subunits in a tRNA import complex was contradicted by further observations (Paris et al. 2009; Cristodero et al. 2010). Recent in vivo investigations by other groups in *T. brucei* led to a different picture. The role of eEF1α in mitochondrial tRNA import was confirmed, while the TIM17 subunit of the protein translocase of the inner membrane (TIM) and the mitochondrial heat-shock protein HSP70 were shown to be required for the process (Tschopp et al. 2011). An additional mitochondrial membrane complex with affinity for tRNAs was purified from *T. brucei*, and the involvement of two membrane proteins in this complex in mitochondrial tRNA import was confirmed (Seidman et al. 2012). One of these proteins is homologous to a component of the *L. tropica* RIC complex mentioned above. Notably, both TIM17 and HSP70, together with HSP20 and HSP60, were found to be associated with the putative *T. brucei* tRNA translocon (Fig. 4.6). Altogether, there is increasing evidence that mitochondrial tRNA import and protein import pathways share components in *Trypanosoma* (Tschopp et al. 2011; Seidman et al. 2012).

4.4.1.2 Mitochondrial Trafficking of Nucleus-Encoded tRNAs in Plants

Plant mitochondria use the universal genetic code and require an extensive collection of tRNA isoacceptors. Depending on the species, one third to one half of these tRNA species are nucleus-encoded and shared with the cytosol (Maréchal-Drouard et al. 1990; Kumar et al. 1996; Glover et al. 2001; Salinas et al. 2008; Duchêne et al. 2009; Sieber et al. 2011; Salinas-Giegé et al. 2015). Substantial differences in the identity of the imported tRNAs have been reported between angiosperms and gymnosperms, dicotyledonous and monocotyledonous plants, but also between closely related species. The mitochondrial genomes of a number of plant species have been completely sequenced (https://www.ncbi.nlm.nih.gov/genome/organelle/), which allowed one to deduce the missing tRNA genes. However, direct experimental analyses are necessary to establish the precise set of imported isoacceptors in a given species, as some tRNA genes present in the mitochondrial genome are not functional. This is the case, for instance, for the $tRNA^{Trp}$ gene found in the *A. thaliana* mtDNA or the $tRNA^{Ser}_{(GCT)}$ gene of sunflower (*Helianthus annuus*) mitochondria. Defective mitochondrial tRNA genes were shown to be compensated through uptake of the corresponding tRNAs from the cytosol (Ceci et al. 1996; Duchêne and Maréchal-Drouard 2001). On the other hand, the reason for the import of some isoacceptors is not necessarily obvious, as it seems to lead to redundancy in decoding. It was proposed, on the contrary, that this redundancy reflects restrictions in codon/anticodon recognition due to the presence of modified nucleotides in the anticodon (Sieber et al. 2011). Other lineages present a contrasting situation. The mitochondrial genome of the bryophyte *Marchantia polymorpha* carries 29 tRNA genes representing 27 different tRNAs, lacking only genes for $tRNA^{Ile}$ decoding the AUU and AUC codons and $tRNA^{Thr}$ decoding the ACA and ACG codons (Oda et al. 1992). Nucleus-encoded $tRNA^{Ile}_{(AAU)}$ and $tRNA^{Thr}_{(AGU)}$ were indeed shown to be present in *M. polymorpha* mitochondria (Akashi et al. 1996, 1997). More puzzling, a priori non-needed cytosolic $tRNA^{Val}_{(AAC)}$ was also recovered in organellar fractions

(Akashi et al. 1998). Conversely, the small mitochondrial genome of the green alga *Chlamydomonas reinhardtii* encodes only 3 tRNAs (Michaelis et al. 1990), while 34 mitochondrial tRNAs are considered to be of nuclear origin. It turned out that the steady-state levels of the imported tRNA isoacceptors are correlated with both the cytosolic and the mitochondrial codon usage (Vinogradova et al. 2009). Moreover, mitochondrial tRNA import and codon usage seem to have coevolved in *C. reinhardtii* (Salinas et al. 2012).

Structural determinants for mitochondrial import in plants have been investigated in transgenic *N. tabacum* cell lines expressing variants of naturally imported or non-imported tRNAs (Delage et al. 2003b; Laforest et al. 2005; Salinas et al. 2005). Mutations or sequence exchanges in the D-domain, the T-domain, or the anticodon inhibited mitochondrial uptake of importable tRNAs. However, introducing features of an importable tRNA into a naturally non-importable tRNA was not sufficient to provide importability. Identity determinants of an importable tRNA are essential, as recognition by the cognate aminoacyl-tRNA synthetase was shown to be necessary for import (Dietrich et al. 1996a). However, recognition by a mitochondrially imported cognate aminoacyl-tRNA synthetase was not sufficient to target a naturally non-imported tRNA to the organelles (Dietrich et al. 1996b). Notably, a U to G mutation at position 34 in the anticodon of the importable $tRNA^{Gly}_{(UCC)}$ inhibited import, in line with the importance of the anticodon for $tRNA^{Gly}$ identity (Giegé et al. 1998). Conversely, the anticodon is not involved in $tRNA^{Ala}$ identity, and introduction of a 4-nucleotide synthetic cargo sequence after position 35 in the anticodon of *A. thaliana* $tRNA^{Ala}_{(UGC)}$ did not impair mitochondrial import in transgenic *N. tabacum* plants (Dietrich et al. 1996a). Similarly, a bean (*Phaseolus vulgaris*) $tRNA^{Leu}$ containing an A > G mutation at position 35 in the anticodon and a 4-nucleotide insertion after position 36 was imported into mitochondria upon expression from a nuclear transgene in transformed potato (*Solanum tuberosum*) plants (Small et al. 1992). Altogether, the data suggest that the aminoacyl-tRNA synthetases are involved in targeting the importable tRNAs to the mitochondrial membrane. The TOM40 and TOM20 components of the protein translocase of the outer membrane (TOM) were reported to contribute to further mitochondrial binding, while the voltage-dependent anion channel (VDAC) might ensure translocation through the outer membrane (Salinas et al. 2006, 2014) (Fig. 4.6). Uptake efficiency and selectivity were shown to require ATP (Delage et al. 2003a). Two additional putative RNA-binding proteins located in the outer membrane of *A. thaliana* mitochondria were recently shown to have a role in tRNA import (Murcha et al. 2016) (Fig. 4.6). Called TRIC1 and TRIC2, for "tRNA import component," they are members of the preprotein and amino acid transporter (PRAT) family characterized by the presence of four transmembrane regions and a conserved domain (Rassow et al. 1999). TRIC1 and TRIC2 interact with components of the TOM and TIM protein translocases and display exposed domains on the mitochondrial outer membrane. They are required for tRNA uptake but not for protein import. From these different data, it was hypothesized that repurposing of a preexisting protein import apparatus provided the tRNA import pathway (Murcha et al. 2016).

4.4.1.3 Mitochondrial Trafficking of Nucleus-Encoded tRNAs in Fungi

Partial localization of the cytosolic $tRNA^{Lys}_{(CUU)}$ to mitochondria in the yeast *Saccharomyces cerevisiae* was also reported very early in the field (Martin et al. 1979) and was subsequently the subject of extensive studies (Tarassov and Entelis 1992; Tarassov et al. 2007). The yeast mtDNA encodes its own $tRNA^{Lys}$, but the imported $tRNA^{Lys}$ seems to have distinct recognition properties for the AAG codons. It thus becomes necessary for adaptation of mitochondrial protein synthesis when position 34 in the anticodon of the mitochondrially encoded $tRNA^{Lys}$ becomes hypomodified at elevated temperatures (Kamenski et al. 2007). Partial mitochondrial localization of cytosolic $tRNA^{Gln}_{UUG}$ and $tRNA^{Gln}_{CUG}$, together with the cognate cytosolic glutaminyl-tRNA synthetase, was also put forward in yeast, implying that a direct aminoacylation pathway generates the Gln-$tRNA^{Gln}$ needed for organellar translation (Rinehart et al. 2005). However, a further debate arose when it was established that the yeast mitochondrion-encoded $tRNA^{Gln}$ is actually mischarged to Glu-$tRNA^{Gln}$ by organelle-localized cytosolic glutamyl-tRNA synthetase and that this intermediate shifts to Gln-$tRNA^{Gln}$ upon transamidation involving a novel type of trimeric tRNA-dependent amidotransferase (Frechin et al. 2009).

Extensive analysis of mutant and variant tRNA versions was developed to address the molecular basis underlying selective mitochondrial import of $tRNA^{Lys}_{(CUU)}$ in *S. cerevisiae* (Entelis et al. 1996, 1998; Kolesnikova et al. 2002). The major structural determinants for import include C34 in the anticodon, as well as the G1:C72 base pair and the U73 position in the aminoacyl acceptor stem. Transfer of these features provided the non-imported cytosolic $tRNA^{Lys}_{(UUU)}$ with the ability to translocate into mitochondria in vitro and in vivo (Entelis et al. 1998). Detailed studies on the targeting and import mechanisms (Brandina et al. 2006; Entelis et al. 2006; Kamenski et al. 2010; Baleva et al. 2015) established that the $tRNA^{Lys}_{(CUU)}$ is initially recruited in the cytosol by an enzyme of the glycolytic pathway, the enolase (Fig. 4.6). Binding to enolase enables a 3% fraction of the pool of aminoacylated $tRNA^{Lys}_{(CUU)}$ to withdraw from cytosolic translation, promoting a drastic conformational rearrangement from the canonical L-shape to a so-called F-form (Kolesnikova et al. 2010). The complex shuttles to the mitochondrial membrane, where the tRNA is transferred to the precursor of the mitochondrial lysyl-tRNA synthetase (Fig. 4.6) and recovers an L-shape. In this respect, a mitochondrion-associated macromolecular complex containing the enolase, further glycolytic enzymes, the precursor of the mitochondrial lysyl-tRNA synthetase, and the $tRNA^{Lys}_{(CUU)}$ was characterized (Brandina et al. 2006). It is hypothesized that ultimately the precursor of the mitochondrial lysyl-tRNA synthetase in some way drives the tRNA into the organelles using the mitochondrial protein import channel (Tarassov et al. 1995a, b) (Fig. 4.6). Regulation of $tRNA^{Lys}_{(CUU)}$ import into yeast mitochondria by the ubiquitin-proteasome system was reported (Brandina et al. 2007).

4.4.1.4 Mitochondrial Trafficking of Nucleus-Encoded tRNAs in Animals

The mitochondrial genome of various metazoa carries only a few tRNA genes (Lithgow and Schneider 2010; Schneider and Maréchal-Drouard 2000). Conversely, mtDNAs of mammals are considered to encode all tRNAs needed to read the adapted genetic code used in the mitochondria of these organisms. Nevertheless, the abovementioned *S. cerevisiae* cytosolic $tRNA^{Lys}_{(CUU)}$ and derivatives thereof could be imported into human mitochondria in vitro (Kolesnikova et al. 2000; Entelis et al. 2001). On the other hand, the mitochondrial $tRNA^{Lys}$ gene appears to be a pseudogene in marsupials, and it was shown that mitochondria from these species contain a $tRNA^{Lys}$ coming from the cytosol (Dörner et al. 2001). Although the mtDNA-encoded $tRNA^{Gln}$ is functional (Nagao et al. 2009), the presence of cytosolic $tRNAs^{Gln}$ in rat and human mitochondria was reported, and isolated human mitochondria were able to take up $tRNAs^{Gln}$ or a bacterial pre-$tRNA^{Asp}$ in vitro (Rubio et al. 2008). Altogether, it appears that mammalian mitochondria have kept the ability to import at least some tRNAs.

Yeast cytosolic $tRNA^{Lys}_{(CUU)}$ derivatives also turned out to be importable into mitochondria in human cell cultures and to partially rescue mitochondrial functions in *trans*-mitochondrial cybrid cells and in patient-derived fibroblasts affected by the $tRNA^{Lys}$ A8344G pathogenic mutation causing the MERRF (myoclonic epilepsy with ragged red fibers) syndrome (Kolesnikova et al. 2004). Applying the same strategy with a recombinant tRNA bearing the identity elements for human mitochondrial leucyl-tRNA synthetase allowed partial rescue the $tRNA^{Leu}_{(UUR)}$ A3243G pathogenic mutation causing the MELAS (mitochondrial encephalomyopathy with lactic acidosis and stroke-like episodes) syndrome (Karicheva et al. 2011).

The mechanism of mitochondrial targeting and uptake of $tRNA^{Lys}$ and derivatives in human cells is likely to resemble that in yeast (see above Sect. 4.4.1.3), as the precursor of the human mitochondrial lysyl-tRNA synthetase promoted the import in vitro and in cultured human cells (Entelis et al. 2001; Gowher et al. 2013). Also, the efficiency of the import into isolated human mitochondria increased in the presence of rabbit or human enolase (Gowher et al. 2013; Baleva et al. 2015). Human enolase binds to yeast $tRNA^{Lys}_{(CUU)}$ and improves tRNA binding to the precursor of the human mitochondrial lysyl-tRNA synthetase (Baleva et al. 2015).

4.4.2 tRNA-Derived Sequences as Mitochondrial RNA Shuttles

Strategies have been developed using signal sequences of naturally imported tRNAs as shuttles to target cargo RNAs into mitochondria in living cells. Extensive analyses with derivatives of the *S. cerevisiae* $tRNA^{Lys}_{CUU}$ led to the design of short importable RNAs containing two domains of the tRNA, the D-arm and a domain that was

called the F-hairpin, joined by a central linker domain (Kolesnikova et al. 2010). Exchanging the central linker of such FD-RNAs for an antigenomic sequence led to repression of the replication of mutated mtDNA carrying a large pathogenic deletion or a point mutation in human cybrid cells (Comte et al. 2013; Tonin et al. 2014).

The D-arm hairpin of the *Leishmania* $tRNA^{Tyr}$ was identified as a determinant for mitochondrial import (Mahapatra et al. 1998; Bhattacharyya et al. 2002) (see above Sect. 4.4.1.1). In an extensive series of experiments, this structure was fused with short antisense RNAs to mitochondrial mRNAs, with antisense DNA oligonucleotides or with long polycistronic RNAs carrying native mtDNA coding sequences. The various chimeric fusions were loaded on isolated RIC complex (see above Sect. 4.4.1.1) and applied to human HepG2 cells and patient-derived cybrids or injected into middle-aged and old rats. In all cases, the expected mitochondrial antisense effect, mtDNA mutation rescue, or stimulation of mitochondrial translation and respiratory capacity was reported (Mukherjee et al. 2008; Mahato et al. 2011; Jash and Adhya 2011; Jash et al. 2012). The D-arm hairpin of the *Leishmania* $tRNA^{Tyr}$ was in turn fused to an antisense RNA oligonucleotide directed against the mitochondrial *COX2* mRNA, so as to promote mitochondrial import. The chimeric transcript was loaded on MITO-Porter transducing/mitochondriotropic nanovesicles (Furukawa et al. 2015). Incubation with HeLa cells resulted in knockdown of the *COX2* target RNA. As the stem-loop structure of the *MRP* RNA (see above Sect. 4.2), the $tRNA^{Tyr}$ D-arm hairpin was grafted to the MITO-Porters as an RNA aptamer, to optimize the nanovesicle functionality (Yamada et al. 2016).

In plants, cargo RNAs were linked as 5′-trailors to a 120-nucleotide tRNA-like sequence taken from the 3′-end of the Turnip yellow mosaic virus (TYMV) genomic RNA (Val et al. 2011). This sequence mimics the mitochondrially importable $tRNA^{Val}$ (Matsuda and Dreher 2004). The chimeric RNAs were expressed from nuclear transgenes and targeted into mitochondria in *N. tabacum* and *A. thaliana*. Using *trans*-cleaving hammerhead ribozymes as cargoes allowed specific knockdown of individual mitochondrial RNAs and characterization of new organellar gene functions (Sultan et al. 2016; Val et al. 2011).

4.4.3 *Mitochondrial Trafficking of the RNase P RNA and tRNA Processing*

The history of RNase P RNA trafficking is somewhat parallel to that of *MRP* RNA localization (see Sect. 4.2). The ubiquitous endoribonuclease RNase P ensures 5′-end processing of tRNA precursors but has further roles and cleaves also substrates such as lncRNAs, rRNAs, and mRNAs (Klemm et al. 2016; Jarrous 2017). After discoveries in the laboratories of S. Altman and T. Cech on ribozymes in the 1970s, it has been considered for decades that throughout prokaryotes and eukaryotes, RNase P is composed of a catalytic RNA responsible for the activity and a protein subunit complement. The concept was readily applied to organelles. Mitochondrial RNase

P activity in yeast was indeed shown to be dependent on protein coded for by nuclear DNA and an RNA subunit encoded by the mtDNA (Miller and Martin 1983; Hollingsworth and Martin 1986; Dang and Martin 1993). The mammalian mtDNA does not encode such an RNA. Nevertheless, Doersen et al. (1985) reported in early studies the enrichment of a HeLa cell mitochondrial RNase P with a nuclease-sensitive activity, implying that the mammalian organellar RNase P would also be a ribonucleoprotein with a catalytic RNA and that the nucleus-encoded catalytic RNA moiety of RNase P would thus be imported into mitochondria (Fig. 4.7). Conflicting further assays failed to detect any nuclear RNase P activity or *H1* RNase P RNA in mitochondrial extracts and implied that mammalian mitochondrial RNase P would not be a ribonucleoprotein (Rossmanith et al. 1995; Rossmanith and Karwan 1998). The nucleus-encoded *H1* RNase P RNA remaining in organellar fractions was then attributed to contamination unrelated to the mitochondrial RNase P activity. However, in agreement with the initial claims (Doersen et al. 1985), a mitochondrion-associated RNase P activity was later extensively purified from HeLa cells and reported to contain an RNA identical in sequence to the *H1* RNA of nuclear RNase P. Upon quantitative evaluation, the study concluded that the levels of *H1* RNA detected in HeLa cell mitochondrial fractions should be adequate to satisfy the mitochondrial tRNA synthesis requirements (Puranam and Attardi 2001). Nevertheless, the mammalian RNase P was finally characterized as an enzyme made only of proteins, through a combination of classical purification and mass spectrometry analyses (Holzmann et al. 2008; Walker and Engelke 2008). Moreover, the activity was reconstituted from the three identified subunits, without any RNA component. The plant mitochondrial RNase P was shown as well to be a protein-only enzyme (Gobert et al. 2010; Gutmann et al. 2012; Lechner et al. 2015).

The question of mitochondrial import of the *H1* RNase P RNA was not closed with these results, and, as in the case of the *MRP* RNA (see above Sect. 4.2), the dispute was revived by the reports focused on the putative role of mitochondrial PNPASE in mitochondrial import of RNAs (Wang et al. 2010, 2012a). The authors proposed that the *H1* RNA actually binds to PNPASE and functions in PNPASE-dependent processing of mitochondrial RNAs (Fig. 4.7). In the frame of these experiments, mitochondrial uptake of human *H1* RNA was tested both in vitro and in vivo. While yeast naturally lacks PNPASE, the uptake was higher with mitochondria isolated from yeast expressing the human PNPASE versus mitochondria from wild-type yeast (Wang et al. 2010). Also, efficient in vivo mitochondrial import of the human *H1* RNA was reported in yeast cells expressing mammalian PNPASE (Wang et al. 2010). Similarly, the *H1* RNA was taken up by mitochondria from wild-type mouse MEF cells or mouse cells expressing the human PNPASE, but not by mitochondria from mutated mouse cell lines lacking PNPASE. As for the *MRP* RNA, *H1* RNA uptake was strongly decreased with liver mitochondria isolated from a mouse with a liver-specific knockout of the PNPASE gene versus mitochondria from wild-type mouse (Wang et al. 2010).

A 20-nucleotide stem-loop structure of the *H1* RNA, similar to that in the *MRP* RNA (see Sect. 4.2), was identified as a signal for PNPASE-dependent organellar uptake (Fig. 4.7), and its fusion to the non-imported *GAPDH* mRNA allowed the

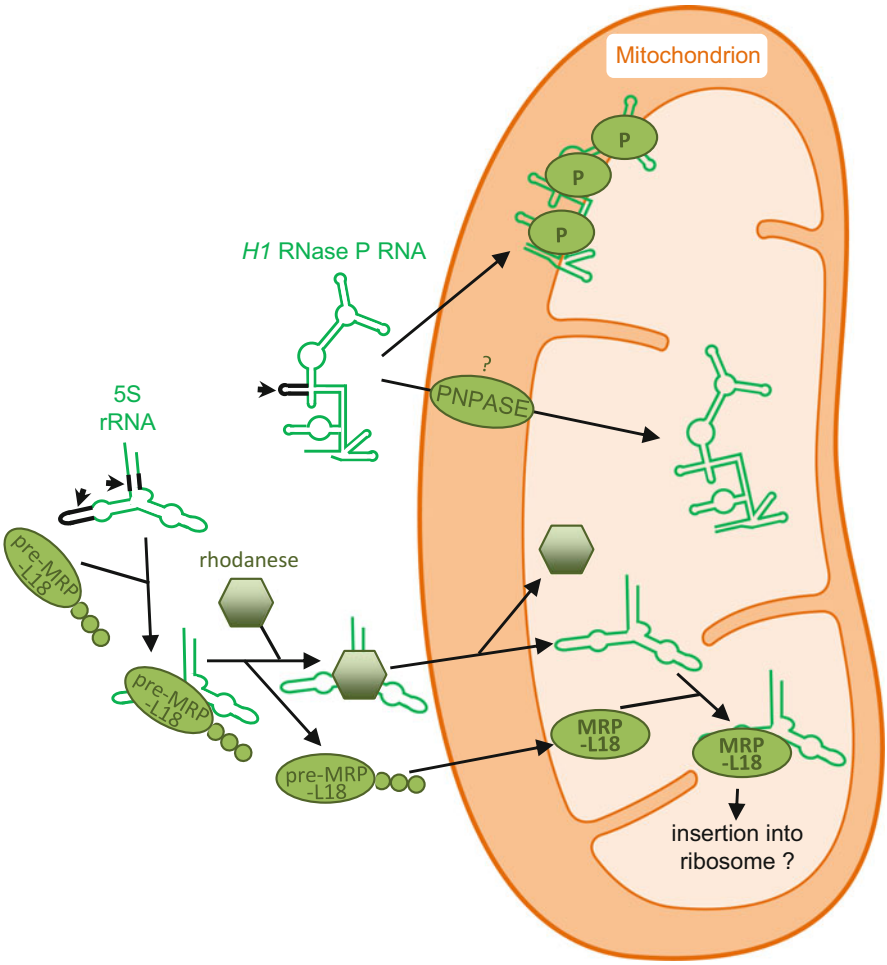

Fig. 4.7 Models for mitochondrial import of RNase P *H1* RNA and 5S rRNA. The regions described as mitochondrial import determinants (Wang et al. 2010; Smirnov et al. 2011) are highlighted as thicker black lines and pointed by arrows in one of the *H1* or 5S rRNA structures. Putative import of the *H1* RNA was proposed to involve PNPASE. Protection of the *H1* RNA against RNase digestion in mitochondrial fractions by tightly bound protein(s) was also implied. The first factor involved in 5S rRNA import is the precursor of the mitochondrial ribosomal protein L18 (pre-MRP-L18) that binds to the Loop D region of the γ-domain in the cytosol. The RNA is subsequently handed over to rhodanese and import proceeds. P represents unknown protein(s)

latter to be taken up into isolated yeast mitochondria (Wang et al. 2010). Human mitochondrial $tRNA^{Trp}$ fused with the *H1* RNA stem-loop structure, but not $tRNA^{Trp}$ alone, was imported into isolated mouse liver mitochondria. A subsequent study reported that the *H1* RNA stem-loop was able to promote PNPASE-dependent uptake of mitochondrial tRNA precursors into isolated mouse or human cybrid mitochondria (Wang et al. 2012b). The imported precursors appeared to be

processed and to function in mitochondrial translation. The strategy was further developed to import *H1* stem-loop/*COX2* mRNA fusion transcripts into mitochondria in human or mouse cells, leading to the production and membrane insertion of the COX2 polypeptide in vivo. Finally, appropriate precursors of wild-type mitochondrial tRNAs were combined with the *H1* stem-loop and with the 3′-UTR sequence that normally targets the mRNA of human mitochondrial ribosomal protein S12 to the organellar surface for translation (Russo et al. 2008). The chimeric RNAs were expressed in human cybrid cell lines and rescued MERRF or MELAS mitochondrial defects due to mutations in the mtDNA affecting tRNA genes (Wang et al. 2012b).

The above data bring evidence that the *H1* RNA is importable into mitochondria. Nevertheless, whether this is a regular process in mammalian cells and whether a protein-only RNase P enzyme and a PNPASE-*H1* RNA processing enzyme coexist in mitochondria, with potentially distinct substrates, is still a matter of contradictory discussion (Wang et al. 2010, 2012a; Rossmanith 2012). The discrepancies in the functions proposed for this RNA in the organelles, together with the low mitochondrial estimates of only 33 to 175 molecules per cell, i.e., still much less than 1 molecule per mitochondrion (Puranam and Attardi 2001; Posakony et al. 1977), account for the main issues of the debate. Especially puzzling in this long debate of conflicting studies are the recent assays of Cannon et al. (2015) mentioned above (Sects. 4.2 and 4.3.3) in which the authors carried out stepwise digitonin-mediated removal of the outer membrane of mouse mitochondria combined with successive RNase treatments. As a result, the *H1* RNA remained resistant to RNase treatment even in the final mitochondrial fraction, i.e., after complete organelle lysis in the presence of the highest digitonin concentration. Nuclease resistance thus appeared to be independent of the integrity of the inner membrane, suggesting that the *H1* RNA was actually protected by tightly bound protein(s) (Fig. 4.7).

4.4.4 Mitochondrial Trafficking of the 5S rRNA

The 5S rRNA is a universal component of ribosomes in prokaryotes and in the cytosol of eukaryotes. A distinct 5S rRNA encoded by the mitochondrial genome is also found in angiosperms and select protist groups, but such a gene is lacking in animal and fungal mitochondria (Valach et al. 2014). Nevertheless, the presence of the cytosolic 5S rRNA in mitochondria was repeatedly reported in mammals (Fig. 4.7). The experimental evidence was based on extensive purification and RNase treatment of mitochondria and mitoplasts and on in vivo expression of a tagged 5S rRNA version (Yoshionari et al. 1994; Magalhaes et al. 1998). The significance and role of this RNA in the organelles initiated one more debate. On the one hand, 5S rRNA was found associated with human mitochondrial ribosomes isolated in mild conditions (Smirnov et al. 2011), while on the other hand, it was not detected when the complete structure of the human mitoribosome was determined at a subnanometer resolution (Amunts et al. 2015; Greber et al. 2015; Englmeier et al. 2017). A third RNA moiety

was indeed present in the solved structures, but it was a $tRNA^{Val}$ or a $tRNA^{Phe}$ encoded by the mitochondrial genome (Amunts et al. 2015; Greber et al. 2015; Rorbach et al. 2016). Whether the contradiction is due to mitoribosome isolation and handling procedures remains to be clarified. It seems unlikely at the present stage that the cytosolic 5S rRNA would be part of the core structure of the mammalian mitochondrial ribosome, which does not necessarily exclude the possibility of circumstantial association. Notably, decrease of 5S rRNA import coincided with a general decrease of mitochondrial translation, pointing to a functional importance of this RNA inside the organelles (Smirnov et al. 2010).

The basis of 5S rRNA subcellular trafficking was further investigated (Smirnov et al. 2008). Extensive in vitro assays with isolated human mitochondria pointed to a putative involvement of the preprotein import channel in the organellar uptake (Entelis et al. 2001). Further in vitro and in vivo assays identified two distinct regions of the human 5S rRNA that appeared to be necessary for mitochondrial targeting: (1) the proximal part of Helix I (in the α-domain) containing the conservative uncompensated G_7-U_{112} pair and (2) the Helix IV/Loop D region (in the γ-domain), which shows several noncanonical structural features (Smirnov et al. 2008) (Fig. 4.7). Conversely, the β-domain turned out to be dispensable for organellar uptake. Mitochondrial targeting of the 5S rRNA was shown to require rhodanese, i.e., the mitochondrial thiosulfate sulfurtransferase, as an import factor (Smirnov et al. 2010) (Fig. 4.7). Evidence was reported that the 5S rRNA, in a particular family A conformation, can bind co-translationally to the nascent form of rhodanese through the above-identified mitochondrial targeting determinants, predominantly the proximal part of Helix I. The RNA is proposed to act as a chaperone, maintaining a misfolded, organellar import-competent conformation of the precursor protein and functionally mimicking the chaperones involved in mitochondrial targeting of preproteins. It was thus speculated that the 5S rRNA might exploit the mitochondrial precursor protein import pathway for its own targeting to mitochondria (Smirnov et al. 2010), a possibility also supported by the above in vitro assays (Entelis et al. 2001). Further studies identified the mitochondrial ribosomal protein L18 (MRP-L18) as a second factor for 5S rRNA organellar targeting (Fig. 4.7). Cytosolic 5S rRNA was reported to bind to human MRP-L18 through the Helix IV/Loop D region also identified above as a mitochondrial targeting determinant (Smirnov et al. 2011). Binding to the precursor of MRP-L18 changes the folding of the 5S rRNA, potentially into the conformation able to recognize the nascent rhodanese. The data altogether led to a working model of 5S rRNA subcellular trafficking (Smirnov et al. 2011). Upon gene transcription in the nucleus, the 5S rRNA is first exported to the cytosol and then redirected to the nucleolus as a complex with the L5 ribosomal protein. In the nucleolus, the RNA and protein are both integrated into ribosomal large subunits. The hypothesis was proposed that pre-MRP-L18 might compete with L5 for 5S rRNA binding in the cytosol and withdraw from the prevailing route to the nucleolus about 1% of the 5S rRNA pool for mitochondrial targeting. Refolded through binding to pre-MRP-L18, the RNA would in turn become a chaperone for nascent rhodanese. Finally, the latter would function as a carrier for mitochondrial uptake (Smirnov et al. 2011) (Fig. 4.7).

Contrasting with this model, the assays of Wang et al. (2010) mentioned above (Sects. 4.2 and 4.4.3) showed that uptake of human 5S rRNA into mitochondria isolated from yeast expressing the human PNPASE was also more efficient when compared to mitochondria from wild-type yeast. Conversely, as for the *MRP* and *H1* RNAs, 5S rRNA uptake into liver mitochondria isolated from a mouse with a liver-specific knockout of the PNPASE gene was severely compromised versus mitochondria from wild-type mouse. Whether these contradictory data reflect distinct translocation pathways or can be integrated remains to be clarified.

Knowledge of import determinants (Smirnov et al. 2008) allowed development of mitochondrial shuttle systems exploiting 5S rRNA trafficking. In vivo mitochondrially importable recombinant 5S rRNAs were obtained by exchanging part of the ß-domain for short foreign cargo sequences of 12 to 14 nucleotides (Smirnov et al. 2008; Comte et al. 2013). Mitochondrial targeting of recombinant 5S rRNA variants carrying antigenomic sequences against the deletion junction in mutant mtDNA causing the Kearns-Sayre syndrome decreased by half the proportion of mutated versus wild-type mtDNA in cultured *trans*-mitochondrial cybrid cells (Comte et al. 2013). Alternatively, a sequence of 21 nucleotides complementary to a region of the mitochondrial *ND5* mRNA or gene was attached as a cargo to the 3′-end of a modified 5S rRNA γ-domain. The chimeric RNA was reported to localize to mitochondria when introduced or expressed in human cells (Zelenka et al. 2014; Zelenka and Jezek 2016).

4.5 Storage of Nucleus-Originating miRNAs and Regulation of the Cytosolic Transcriptome

Hundreds of nucleus-encoded miRNAs were found associated with mitochondria upon high-throughput or specific analyses, with substantial differences between cell types (Bandiera et al. 2011; Barrey et al. 2011; Bian et al. 2010; Das et al. 2012; Dasgupta et al. 2015; Kren et al. 2009; Jagannathan et al. 2015; Mercer et al. 2011; Sripada et al. 2012a; Vargas et al. 2016; Wang et al. 2015; Wang and Springer 2015). Only a limited set of these miRNAs was shown or predicted to have mtDNA-encoded targets (see above Sect. 4.3.1), while the vast majority is still likely to target nucleus-encoded mRNAs. The data as a whole widen the view of mitochondrial contribution to cellular homeostasis but raise many questions in terms of RNA trafficking, silencing mechanisms, and regulation pathways (Bandiera et al. 2013; Bian et al. 2010; Kren et al. 2009; Leung 2015; Sripada et al. 2012a; Wang and Springer 2015; Wang et al. 2015; Zhang et al. 2014). In terms of mitochondrial localization, at least three open possibilities still have to be considered. Firstly, rather than inside, some miRNAs seemed to localize at the surface of mitochondria, combined with Ago2 or Ago3 and their target mRNAs. It was thus suggested that the mitochondrial outer membrane might provide a platform on which to assemble the miRNA/RISC complexes in order to regulate subcellular site-specific protein

levels (Sripada et al. 2012a). A large subset of cytosolic mRNAs in various organisms actively localizes to the mitochondrial surface for translation (Lesnik et al. 2015; Vincent et al. 2017) and might indeed become there a target for miRNA-mediated silencing. Nevertheless, the vast majority of the studies report the presence of nucleus-encoded miRNAs inside mitochondria, and the concept has emerged that the organelles might constitute a "reservoir" or a "warehouse" for storage and "on demand" release of miRNAs for regulation processes or stress response (Bandiera et al. 2013; Kren et al. 2009; Wang and Springer 2015). Notably, the reports essentially rely on RNA preparations from RNase-treated organelles with an intact outer membrane. A more rarely considered second open possibility would thus be the localization of miRNAs in the mitochondrial intermembrane space, rather than in the matrix. To target cytosolic mRNAs, nucleus-originating miRNAs localizing in mitochondria need to translocate both into (as discussed above in Sect. 4.3.1) and out of the organelles. In our own experiments, the inner membrane was a much stronger barrier than the outer membrane for translocation of longer RNAs, and it cannot be excluded that miRNA exchange would be easier between the cytosol and the intermembrane space. miR-146a, miR-103, and miR-16 were indeed found to be enriched in the intermembrane space of human mitochondria (Mercer et al. 2011). With respect to these first two possibilities, it was noted that some miRNAs or miRNA targets were depleted in mitoplast extracts, i.e., after elimination of the outer membrane, versus fractions from intact mitochondria (Mercer et al. 2011; Sripada et al. 2012a). The import of nucleus-originating miRNAs into the mitochondrial matrix compartment and their subsequent release remains the third possibility and the most challenging concept. In earlier experiments, it was suggested that human mitochondria can export tRNAs, as mitochondrial $tRNA^{Lys}$ and $tRNA^{Met}$ were immunoprecipitated with Ago2 from cytosolic extracts (Beitzinger et al. 2007; Maniataki and Mourelatos 2005), but the existence of a mitochondrial RNA export pathway has not been investigated per se.

While the trafficking mechanisms remain to be clarified, data have emerged that indeed point to a role of mitochondria in regulation pathways driven by nucleus-encoded miRNAs and to a further layer of cross-talk between mitochondria, the nucleus, and the rest of the cell (Bandiera et al. 2011; Leung 2015; Wang and Springer 2015). The set of miRNAs associated with hippocampal mitochondria changed following a controlled cortical impact injury in rats and might be involved in the regulation of the response to traumatic brain injury (Wang et al. 2015). Similarly, the profile of miRNAs associated with mitochondria was altered in diabetic mice (Bian et al. 2010). It was shown as well that a given nucleus-encoded miRNA can have targets both in the cytosol and in the organelles. Remarkably, miR-1 showed opposite effects in the two compartments upon transfection into mouse myoblasts, repressing translation of its cytosolic target RNAs *HDAC4* and *ELL2* while increasing translation of the mitochondrial *ND1* and *COX1* mRNAs (Zhang et al. 2014). Bioinformatic screenings predicted that mitochondrion-associated nucleus-originating miRNAs might have mRNA targets directly related to organelle biogenesis or organelle-specific functions but also targets involved in general processes like nucleotide metabolism, cell cycle, apoptosis, cell

proliferation, and differentiation (Bandiera et al. 2011; Bian et al. 2010; Kren et al. 2009; Sripada et al. 2012a; Wang and Springer 2015). The significance of miRNA localization to mitochondria in the general organization and regulation of the cell has thus been further discussed (Leung 2015). It was hypothesized that mitochondria, as organelles able to integrate RNAs, might serve as vehicles to deliver miRNAs to other cellular compartments (Wang and Springer 2015). Interactions of compartments like the P-bodies or the endoplasmic reticulum with the dynamic mitochondrial network of mammalian cells might support exchange of miRNAs for specialized or localized gene regulation. Mitochondria indeed establish dynamic contacts with P-bodies, i.e., cytoplasmic granules related to mRNA turnover and RNA interference, and inactivation of mitochondria strongly impairs miRNA-mediated RNAi (Huang et al. 2011).

4.6 Regulation of the Nuclear Transcriptome Through Mitochondrially Generated ncRNAs

Conversely to nucleus-originating ncRNAs regulating gene expression in mitochondria, data is emerging that suggest potential nuclear regulation mediated by mtDNA-derived transcripts (Dong et al. 2017; Pozzi et al. 2017).

4.6.1 Mitochondrially Generated Small ncRNAs

The possibility that small ncRNAs generated from the mitochondrial genome might take part in retrograde signaling and influence nuclear gene expression was tentatively explored in the Manila clam *Ruditapes philippinarum* (Pozzi et al. 2017). This organism shows unusual mitochondrial inheritance, i.e., doubly uniparental inheritance (DUI). Two mitochondrial lineages are found in gametes, one transmitted through eggs (female-inherited) and the other through sperm (male-inherited), with up to 43% nucleotide divergence between the mtDNAs (Doucet-Beaupré et al. 2010). Considering, as discussed above, that there are many conceivable ways in which the mitochondrial genome has the potential to generate regulatory RNAs, Pozzi et al. (2017) investigated whether mtDNA-derived small ncRNAs might influence germ-line gene expression and fate in *R. philippinarum*. Sequencing of small RNA libraries derived from enriched mitochondria isolated from *R. philippinarum* gonad samples of both sexes provided a selection of 14 putative small ncRNAs (18–35 nucleotides) mapping essentially to unassigned regions and tRNA genes in the 22 kb mtDNA that contains intergenic sequences (Pozzi et al. 2017; Hwang et al. 2016). At least one target nuclear gene was predicted for each of these candidates, while no target gene was identified in the mitochondrial genome. The predicted nuclear targets take part in a wide range of biological processes like

microtubule dynamics and chromatin remodeling but also include the NR0B1 protein that is crucial for sex determination. Although the presence of NUMTs has not been reported so far for molluscs, as in other studies it could not be excluded that the candidate small ncRNAs might derive from mitochondrial sequences integrated into the nuclear genome (Pozzi et al. 2017). This work again raises the question of RNA migration outside of the organelles. Release of mitochondrial material into the germplasm of *R. philippinarum* was described (Milani et al. 2011) and would be as well a reasonable mechanism for recovering ncRNAs in the cytosol (Pozzi et al. 2017).

4.6.2 Mitochondrially Generated lncRNAs

A set of chimeric, mtDNA-mapping lncRNAs with surprising compositions was identified in mouse and human, leading to an abundant series of reports. The transcripts (1.6 to 2.4 kb) were chimeras combining the mitochondrial 16S rRNA with a 5′-leader sequence derived from the complementary strand (SncmtRNA, SncmtRNA-2) or the antisense mitochondrial 16S rRNA with a 5′-leader sequence derived from the sense strand (ASncmtRNA-1, ASncmtRNA-2) (Burzio et al. 2009; Villegas et al. 2000, 2002b, 2007; Villota et al. 2012) (Fig. 4.8). Mitochondrial transcription was required for the synthesis of these putative lncRNAs, and the sequence was 100% identical to the mtDNA. In the absence of a suitable mitochondrial transcription template, it was proposed that these RNAs might result from posttranscriptional *trans*-splicing in the organelles (Villegas et al. 2000, 2007). Nuclear localization was repeatedly reported (Landerer et al. 2011; Villegas et al. 2000, 2002a), raising the question of the mechanisms able to ensure a putative trafficking of such large and structured RNAs out of the organelles (Fig. 4.8). The presence of the chimeric transcripts in both the mitochondria and the nucleus suggested a role in mitochondrial-nuclear communication and retrograde signaling (Dong et al. 2017; Landerer et al. 2011).

In the nucleus, the SncmtRNAs appeared to associate with the heterochromatin and the nucleoli (Fig. 4.8). They were present in both normal and cancer cells and were proposed to be involved in cell proliferation and cell cycle regulation (Villegas et al. 2002b, 2007). Conversely, the expression of the ASncmtRNAs was downregulated in tumor cell lines and tumors, suggesting that they might function as mitochondrion-encoded tumor suppressors (Burzio et al. 2009; Landerer et al. 2011; Rivas et al. 2012). While not affecting the viability of normal cells, knockdown of the ASncmtRNAs with antisense oligonucleotides inhibited proliferation and induced caspase-dependent apoptosis in tumor cell lines, likely through miRNA interference of survivin expression (Vidaurre et al. 2014). ASncmtRNA knockdown was shown to inhibit tumor growth and metastasis in murine models (Borgna et al. 2017; Lobos-Gonzalez et al. 2016). The ASncmtRNAs were proposed to be a vulnerability or "Achilles' heel" of cancer cells and to constitute potential targets for cancer therapy (Owen 2017; Vidaurre et al. 2014). They also seemed to be

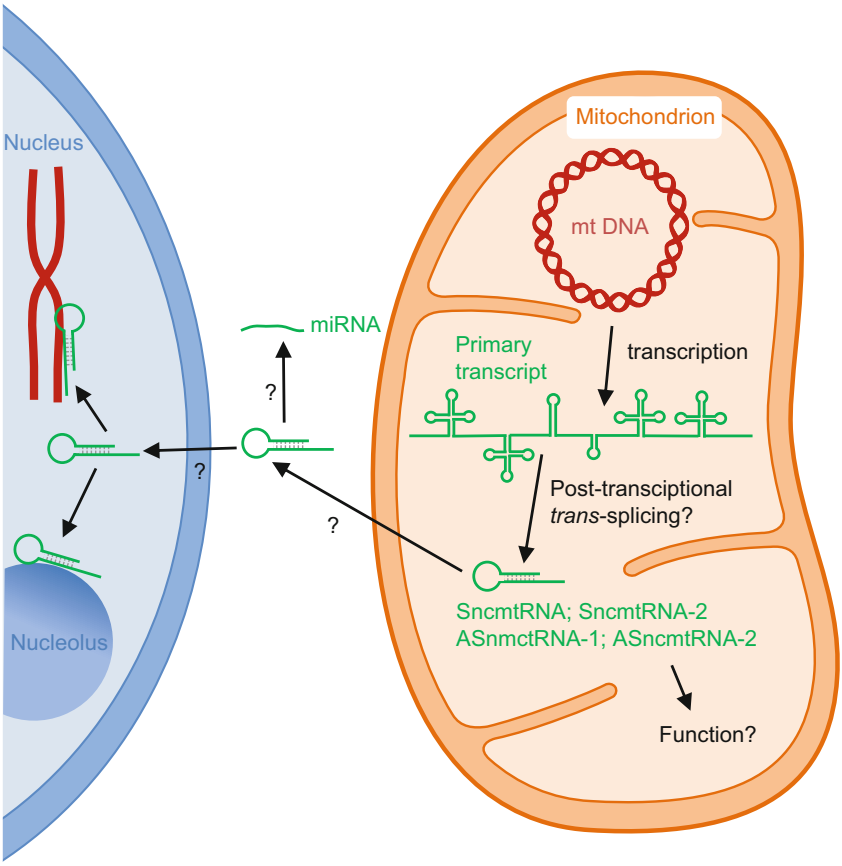

Fig. 4.8 Formation of hairpin lncRNAs in mitochondria and further localization. Processing of mitochondrial primary transcripts followed by putative *trans*-splicing was proposed to generate chimeric hairpin lncRNAs based on sense and antisense sequences from the mitochondrial 16S rRNA. These are subsequently recovered in the nucleus, associated with both the nucleolus and to heterochromatin. Alternatively, they might generate miRNAs in the cytosol

associated with miRNA-mediated aging processes (Bianchessi et al. 2015). ASncmtRNA-2 was upregulated in diabetic kidney (Gao et al. 2017).

Based on their structure, on bioinformatic simulation and on co-immunoprecipitation with Dicer, it was speculated that ASncmtRNAs might themselves give rise to miRNAs (Vidaurre et al. 2014) (Fig. 4.8). In particular, ASncmtRNA-2 was hypothesized to be a mitochondrial noncanonical precursor of hsa-miR-4485 and hsa-miR-1973 and to regulate the cell cycle through miRNA production during replicative senescence (Bianchessi et al. 2015). On the other hand, a 63-nucleotide 5′-fragment released upon processing of SncmtRNA (Villegas et al. 2007) might trap the hsa-miR-620 miRNA and thus modulate the regulation of the

corresponding target genes (Villota et al. 2012). In the study by Sripada et al. (2017) (Sect. 4.3.1), mtDNA depletion but also inhibition of nuclear transport or Dicer decreased the level of miR-4485 in mitochondria, with an enrichment of the precursor in the cytosol. Although they did not rule out the possibility that a subfraction of miR-4485 originates from the mitochondrial genome, the authors concluded that this miRNA is likely encoded by the nuclear genome, processed through the canonical miRNA pathway, and translocated to mitochondria.

4.7 Circulating Mitochondrially Generated ncRNAs

Extensive analyses have established that ncRNAs are critical contributors to cardiovascular pathophysiology. A number of them turned out to circulate in the body fluids and were highlighted as biomarkers of cardiovascular pathologies (Viereck and Thum 2017). Remarkably, among these ncRNAs some actually had a putative mitochondrial origin. Survivors from heart attack often develop cardiac remodeling and heart failure. Global transcriptomic analyses in plasma RNA from patients with or without left ventricular remodeling after myocardial infarction identified a set of seven highly abundant lncRNAs mapping to the mtDNA and consistently amplified throughout the series of plasma samples (Kumarswamy et al. 2014). Their levels were correlated positively with each other. Two of them, lncRNAs uc004cos.4 and uc022bqs, predicted future cardiac remodeling in patients and were significantly downregulated during the early stage of the process. The uc022bqs.1 lncRNA had the greatest association with left ventricular remodeling and was called LIPCAR for "Long Intergenic noncoding RNA Predicting CARdiac remodeling." LIPCAR levels increased at late stages of post-myocardial infarction remodeling and appeared to be a prognostic indicator for chronic heart failure (Kumarswamy et al. 2014). The same set of seven mtDNA-mapping lncRNAs was amplified in serum of patients with hypertrophic obstructive cardiomyopathy or with hypertrophic nonobstructive cardiomyopathy and correlated with clinical parameters (Kitow et al. 2016). In this set, the uc004cov.4 and uc022bqu.1 putative mitochondrial lncRNAs were able to identify hypertrophic obstructive cardiomyopathy as useful clinical biomarkers. LIPCAR was also inversely associated with diastolic function in patients with type 2 diabetes (de Gonzalo-Calvo et al. 2016) and significantly increased in patients with coronary artery disease (Zhang et al. 2017).

In terms of structure, similarly to the abovementioned SncmtRNA and ASncmtRNA lncRNAs (Sect. 4.6.2), LIPCAR appeared to be a chimera, including in this case (−) strand sequences mapping to the 5′ part of the *COX2* gene and the 3′ part of the *CYTb* gene from the human mitochondrial genome. The sequence roughly assembles two of the lncRNAs detected by Yang et al. (Sect. 4.3.4) (Kumarswamy et al. 2014; Yang et al. 2014). The detection of lncRNAs with a putative mitochondrial origin circulating in body fluids adds a further layer of question marks (Dorn 2014). What is the meaning of such RNAs? If they are really assembled from mitochondrial transcripts and do not originate from NUMTs (Richly and Leister 2004; Woischnik and Moraes 2002), do

they have an initial function in the organelles? Are they able to exit the organelles through a physiological pathway for regulation purposes, or do they simply escape the scavenging processes associated with mitophagy? Similarly, one can wonder whether they can be excreted from the cells or whether they are released upon cell damage.

4.8 Conclusion

As detailed throughout this chapter, a plethora of nucleus-encoded ncRNAs has been reported to localize to/into mitochondria, to be stored there and released when required, to be handed over from there to further compartments like P-bodies and the endoplasmic reticulum, or to function inside the organelles in replication, translation, or regulation. Reciprocally, it has been put forward that mitochondria themselves produce ncRNAs to control their own gene expression but also to regulate nuclear gene expression and potentially ensure systemic functions through body fluids. The reports altogether draw a complex and integrated network of RNA trafficking and intercompartment communication. In particular, regulatory RNAs and not only proteins seem to be sent to mitochondria as an anterograde response, while retrograde signaling would involve transport of mitochondrial transcripts. This exciting picture, however, still suffers from weaknesses. The never-ending controversies reported here, for instance, in the case of the *MRP* and *H1* RNAs (Sects. 4.2 and 4.4.3), or the observations that some RNAs can remain resistant to RNase in final mitochondrial lysates independently of inner membrane integrity (Sects. 4.3.3 and 4.4.3) (Cannon et al. 2015) illustrate how difficult it is still to convincingly establish whether a given nucleus-encoded RNA is present or not in the inner organellar compartment. Conversely, it is often difficult as well to definitely conclude that an RNA mapping to the mitochondrial genome is synthesized in the organelles. In the absence of functional pressure, pieces of mtDNA integrated into the nuclear genome (NUMTs) (Richly and Leister 2004; Woischnik and Moraes 2002) tend to undergo sequence drift, which can help to distinguish them from authentic mitochondrial genes. However, the criterion is not considered in many studies, it is not obvious for smaller RNAs, and the original sequence might remain for recent transfers. Although likely to be rare, expression of NUMTs can occur (Song et al. 2013). Thus, following massive identification of candidate mitochondrial ncRNAs, nucleus-encoded or mtDNA-derived, and prediction of putative targets, the time has come to clear the field by advancing robust evidence for relevant functions of individual candidates in mitochondria. The functional data reported for miR-181c, miR-378, miR-21, miR-1, or miR-4485 (Sect. 4.3.1) point in that direction.

A further challenging concept is the extensive multidirectional intercompartment trafficking that is implied by the reported ncRNA distribution. A better knowledge of the paths that allow RNA translocation through organellar membranes would potentially render the proposed scheme more acceptable. Unfortunately, as illustrated throughout this chapter, several decades of research have brought only limited consensual understanding of the molecular mechanisms supporting RNA transport

into mitochondria, and the existence of a reverse mechanism that would promote RNA export from the organelles has not been investigated. Widening the debate, but even more puzzling, mitochondria are also competent for active DNA uptake, at least in vitro (Koulintchenko et al. 2003, 2006). Double-stranded constructs of over 11 kb could be taken up into isolated plant organelles (Ibrahim et al. 2011), and the incorporated DNA was functional *in organello* for transcription, recombination, and repair (Boesch et al. 2009, 2010; Koulintchenko et al. 2003, 2006; Mileshina et al. 2011). DNA export from mitochondria has been tested, but it remains difficult to ascertain, as non-specific membrane leakage or damage can always be objected to such data (Klimenko et al. 2011; Patrushev et al. 2004, 2006). RNA and DNA uptake mechanisms are likely to share common features (Weber-Lotfi et al. 2015). While the principle of nucleic acid trafficking into mitochondria is now widely admitted, the mechanistic picture that emerges from the currently available data remains complex and multifaceted (Campo et al. 2017; Konstantinov et al. 2016). Several routes are likely to be open, differing between organisms or coexisting in the same cell. Whether these are all pathways specifically established for nucleic acids or whether some reflect opportunistic hijacking of regular translocation routes of other components, like the protein import pathways, or taking advantage of molecular assemblies, like the permeability transition pore, is still an open discussion (Weber-Lotfi et al. 2015).

Finally, the observation of long chimeric ncRNAs made of mitochondrial sequences in the nucleo-cytosolic compartment (SncmtRNA, SncmtRNA-2, ASncmtRNA-1, ASncmtRNA-2; Sect. 4.6.2) or in body fluids (LIPCAR; Sect. 4.7) points to a further intriguing field that combines the above two concepts. If such RNAs are indeed synthesized in mitochondria, this implies that there is a route to release long RNAs from the organelles. On the other hand, the chimeras cannot be synthesized as such from the mtDNA and need to be assembled from separate transcripts. *Trans*-splicing was put forward as a mechanism (Villegas et al. 2000, 2007), but in the absence of introns, one might wonder whether there is an appropriate splicing machinery in mammalian mitochondria. An alternative idea that cannot be excluded would be that chimeric RNAs are synthesized from NUMTs, directly from chimeric mitochondrial pseudogenes in the nuclear genome, or upon nuclear *trans*-splicing of distinct transcripts. Human cells are considered to express numerous chimeric RNAs, for the vast majority through unknown processes. Recent bioinformatic mining of available EST databases pointed out fusions of nuclear sequences but also fusions of mitochondrial sequences and even fusions between mitochondrial sequences and nuclear sequences (Yang et al. 2013). In combination with a putative mitochondrial in and out trafficking of transcripts, a full repertoire of possibilities is thus open to speculation for the formation and spread of chimeric lncRNAs including mitochondrial sequences.

Acknowledgments We acknowledge the support of our work by the French Centre National de la Recherche Scientifique (CNRS, UPR2357), the University of Strasbourg, the Agence Nationale de la Recherche (grant numbers ANR-06-MRAR-037-02, ANR-09-BLAN-0240-01), and the Ministère de la Recherche et de l'Enseignement Supérieur (Investissements d'Avenir/Laboratoire d'Excellence MitoCross, grant number ANR-11-LABX-0057_MITOCROSS).

References

Akashi K, Sakurai K, Hirayama J, Fukuzawa H, Ohyama K (1996) Occurrence of nuclear-encoded $tRNA^{Ile}$ in mitochondria of the liverwort *Marchantia polymorpha*. Curr Genet 30:181–185

Akashi K, Hirayama J, Takenaka M, Yamaoka S, Suyama Y, Fukuzawa H, Ohyama K (1997) Accumulation of nuclear-encoded $tRNA^{Thr}$ (AGU) in mitochondria of the liverwort *Marchantia polymorpha*. Biochim Biophys Acta 1350:262–266

Akashi K, Takenaka M, Yamaoka S, Suyama Y, Fukuzawa H, Ohyama K (1998) Coexistence of nuclear DNA-encoded $tRNA^{Val}$(AAC) and mitochondrial DNA-encoded $tRNA^{Val}$(UAC) in mitochondria of a liverwort *Marchantia polymorpha*. Nucleic Acids Res 26:2168–2172

Amunts A, Brown A, Toots J, Scheres SHW, Ramakrishnan V (2015) Ribosome the structure of the human mitochondrial ribosome. Science 348:95–98

Asin-Cayuela J, Gustafsson CM (2007) Mitochondrial transcription and its regulation in mammalian cells. Trends Biochem Sci 32:111–117

Baev V, Milev I, Naydenov M, Vachev T, Apostolova E, Mehterov N, Gozmanva M, Minkov G, Sablok G, Yahubyan G (2014) Insight into small RNA abundance and expression in high- and low-temperature stress response using deep sequencing in *Arabidopsis*. Plant Physiol Biochem 84:105–114

Baleva M, Gowher A, Kamenski P, Tarassov I, Entelis N, Masquida B (2015) A moonlighting human protein is involved in mitochondrial import of tRNA. Int J Mol Sci 16:9354–9367

Bandiera S, Ruberg S, Girard M, Cagnard N, Hanein S, Chretien D, Munnich A, Lyonnet S, Henrion-Caude A (2011) Nuclear outsourcing of RNA interference components to human mitochondria. PLoS One 6:e20746

Bandiera S, Mategot R, Girard M, Demongeot J, Henrion-Caude A (2013) MitomiRs delineating the intracellular localization of microRNAs at mitochondria. Free Radic Biol Med 64:12–19

Baradan R, Hollander JM, Das S (2017) Mitochondrial miRNAs in diabetes: just the tip of the iceberg. Can J Physiol Pharmacol 95:1156–1162

Barman B, Bhattacharyya SN (2015) mRNA targeting to endoplasmic reticulum precedes ago protein interaction and microRNA (miRNA)-mediated translation repression in mammalian cells. J Biol Chem 290:24650–24656

Barrey E, Saint-Auret G, Bonnamy B, Damas D, Boyer O, Gidrol X (2011) Pre-microRNA and mature microRNA in human mitochondria. PLoS One 6:e20220

Beitzinger M, Peters L, Zhu JY, Kremmer E, Meister G (2007) Identification of human microRNA targets from isolated argonaute protein complexes. RNA Biol 4:76–84

Bennett JL, Clayton DA (1990) Efficient site-specific cleavage by RNase MRP requires interaction with two evolutionarily conserved mitochondrial RNA sequences. Mol Cell Biol 10:2191–2201

Bhattacharyya SN, Chatterjee S, Adhya S (2002) Mitochondrial RNA import in *Leishmania tropica*: aptamers homologous to multiple tRNA domains that interact cooperatively or antagonistically at the inner membrane. Mol Cell Biol 22:4372–4382

Bhattacharyya SN, Chatterjee S, Goswami S, Tripathi G, Dey SN, Adhya S (2003) "Ping-pong" interactions between mitochondrial tRNA import receptors within a multiprotein complex. Mol Cell Biol 23:5217–5224

Bian Z, Li LM, Tang R, Hou DX, Chen X, Zhang CY, Zen K (2010) Identification of mouse liver mitochondria-associated miRNAs and their potential biological functions. Cell Res 20:1076–1078

Bianchesssi V, Badi I, Bertolotti M, Nigro P, D'Alessandra Y, Capogrossi MC, Zanobini M, Pompilio G, Raucci A, Lauri A (2015) The mitochondrial lncRNA ASncmtRNA-2 is induced in aging and replicative senescence in endothelial cells. J Mol Cell Cardiol 81:62–70

Boesch P, Ibrahim N, Paulus F, Cosset A, Tarasenko V, Dietrich A (2009) Plant mitochondria possess a short-patch base excision DNA repair pathway. Nucleic Acids Res 37:5690–5700

Boesch P, Ibrahim N, Dietrich A, Lightowlers RN (2010) Membrane association of mitochondrial DNA facilitates base excision repair in mammalian mitochondria. Nucleic Acids Res 38:1478–1488

Bonen L (2008) *Cis*- and *trans*-splicing of group II introns in plant mitochondria. Mitochondrion 8:26–34

Borgna V, Villegas J, Burzio VA, Belmar S, Araya M, Jeldes E, Lobos-Gonzalez L, Silva V, Villota C, Oliveira-Cruz L, Lopez C, Socias T, Castillo O, Burzio LO (2017) Mitochondrial ASncmtRNA-1 and ASncmtRNA-2 as potent targets to inhibit tumor growth and metastasis in the RenCa murine renal adenocarcinoma model. Oncotarget 8:43692–43708

Borralho PM, Rodrigues CM, Steer CJ (2015) microRNAs in mitochondria: an unexplored niche. Adv Exp Med Biol 887:31–51

Bouzaidi-Tiali N, Aeby E, Charriere F, Pusnik M, Schneider A (2007) Elongation factor 1a mediates the specificity of mitochondrial tRNA import in *T. brucei*. EMBO J 26:4302–4312

Brandina I, Graham J, Lemaitre-Guillier C, Entelis N, Krasheninnikov I, Sweetlove L, Tarassov I, Martin RP (2006) Enolase takes part in a macromolecular complex associated to mitochondria in yeast. Biochim Biophys Acta 1757:1217–1228

Brandina I, Smirnov A, Kolesnikova O, Entelis N, Krasheninnikov IA, Martin RP, Tarassov I (2007) tRNA import into yeast mitochondria is regulated by the ubiquitin-proteasome system. FEBS Lett 581:4248–4254

Breving K, Esquela-Kerscher A (2010) The complexities of microRNA regulation: mirandering around the rules. Int J Biochem Cell Biol 42:1316–1329

Brown GG, Colas des Francs-Small C, Ostersetzer-Biran O (2014) Group II intron splicing factors in plant mitochondria. Front Plant Sci 5:35

Bunch H (2017) Gene regulation of mammalian long non-coding RNA. Mol Gen Genomics 293:1–15. https://doi.org/10.1007/s00438-00017-01370-00439

Burger G, Zhu Y, Littlejohn TG, Greenwood SJ, Schnare MN, Lang BF, Gray MW (2000) Complete sequence of the mitochondrial genome of *Tetrahymena pyriformis* and comparison with *Paramecium aurelia* mitochondrial DNA. J Mol Biol 297:365–380

Burger G, Gray MW, Forget L, Lang BF (2013) Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists. Genome Biol Evol 5:418–438

Burzio VA, Villota C, Villegas J, Landerer E, Boccardo E, Villa LL, Martinez R, Lopez C, Gaete F, Toro V, Rodriguez X, Burzio LO (2009) Expression of a family of noncoding mitochondrial RNAs distinguishes normal from cancer cells. Proc Natl Acad Sci U S A 106:9430–9434

Buxbaum AR, Haimovich G, Singer RH (2015) In the right place at the right time: visualizing and understanding mRNA localization. Nat Rev Mol Cell Biol 16:95–109

Campbell CT, Kolesar JE, Kaufman BA (2012) Mitochondrial transcription factor a regulates mitochondrial transcription initiation, DNA packaging, and genome copy number. Biochim Biophys Acta 1819:921–929

Campo ML, Peixoto PM, Martinez-Caballero S (2017) Revisiting trends on mitochondrial megachannels for the import of proteins and nucleic acids. J Bioenerg Biomembr 49:75–99

Cannon MV, Irwin MH, Pinkert CA (2015) Mitochondrially-imported RNA in drug discovery. Drug Dev Res 76:61–71

Carrer M, Liu N, Grueter CE, Williams AH, Frisard MI, Hulver MW, Bassel-Duby R, Olson EN (2012) Control of mitochondrial metabolism and systemic energy homeostasis by microRNAs 378 and 378. Proc Natl Acad Sci U S A 109:15330–15335

Ceci LR, Veronico P, Gallerani R (1996) Identification and mapping of tRNA genes on the *Helianthus annuus* mitochondrial genome. DNA Seq 6:159–166

Chang DD, Clayton DA (1987) A mammalian mitochondrial RNA processing activity contains nucleus-encoded RNA. Science 235:1178–1184

Chang DD, Clayton DA (1989) Mouse RNAase MRP RNA is encoded by a nuclear gene and contains a decamer sequence complementary to a conserved region of mitochondrial RNA substrate. Cell 56:131–139

Charriere F, Helgadottir S, Horn EK, Soll D, Schneider A (2006) Dual targeting of a single tRNA (Trp) requires two different tryptophanyl-tRNA synthetases in *Trypanosoma brucei*. Proc Natl Acad Sci U S A 103:6847–6852

Chatterjee S, Home P, Mukherjee S, Mahata B, Goswami S, Dhar G, Adhya S (2006) An RNA-binding respiratory component mediates import of type II tRNAs into *Leishmania* mitochondria. J Biol Chem 281:25270–25277

Chen LL (2016) Linking long noncoding RNA localization and function. Trends Biochem Sci 41:761–772

Chin A, Lecuyer E (2017) RNA localization: making its way to the center stage. Biochim Biophys Acta 1861:2956–2970

Chiu N, Chiu A, Suyama Y (1975) Native and imported transfer RNA in mitochondria. J Mol Biol 99:37–50

Colley SM, Leedman PJ (2009) SRA and its binding partners: an expanding role for RNA-binding coregulators in nuclear receptor-mediated gene regulation. Crit Rev Biochem Mol Biol 44:25–33

Comte C, Tonin Y, Heckel-Mager AM, Boucheham A, Smirnov A, Aure K, Lombes A, Martin RP, Entelis N, Tarassov I (2013) Mitochondrial targeting of recombinant RNAs modulates the level of a heteroplasmic mutation in human mitochondrial DNA associated with Kearns Sayre syndrome. Nucleic Acids Res 41:418–433

Crausaz Esseiva A, Maréchal-Drouard L, Cosset A, Schneider A (2004) The T-stem determines the cytosolic or mitochondrial localization of trypanosomal $tRNAs^{Met}$. Mol Biol Cell 15:2750–2757

Cristodero M, Seebeck T, Schneider A (2010) Mitochondrial translation is essential in bloodstream forms of *Trypanosoma brucei*. Mol Microbiol 78:757–769

Dang YL, Martin NC (1993) Yeast mitochondrial RNase P. Sequence of the RPM2 gene and demonstration that its product is a protein subunit of the enzyme. J Biol Chem 268:19791–19796

Das S, Ferlito M, Kent OA, Fox-Talbot K, Wang R, Liu D, Raghavachari N, Yang Y, Wheelan SJ, Murphy E, Steenbergen C (2012) Nuclear miRNA regulates the mitochondrial genome in the heart. Circ Res 110:1596–1603

Das S, Bedja D, Campbell N, Dunkerly B, Chenna V, Maitra A, Steenbergen C (2014) miR-181c regulates the mitochondrial genome, bioenergetics, and propensity for heart failure *in vivo*. PLoS One 9:e96820

Dasgupta N, Peng Y, Tan Z, Ciraolo G, Wang D, Li R (2015) miRNAs in mtDNA-less cell mitochondria. Cell Death Discov 1:15004

de Gonzalo-Calvo D, Kenneweg F, Bang C, Toro R, van der Meer RW, Rijzewijk LJ, Smit JW, Lamb HJ, Llorente-Cortes V, Thum T (2016) Circulating long-non coding RNAs as biomarkers of left ventricular diastolic function and remodelling in patients with well-controlled type 2 diabetes. Sci Rep 6:37354

Delage L, Dietrich A, Cosset A, Maréchal-Drouard L (2003a) *In vitro* import of a nuclearly encoded tRNA into mitochondria of *Solanum tuberosum*. Mol Cell Biol 23:4000–4012

Delage L, Duchêne AM, Zaepfel M, Maréchal-Drouard L (2003b) The anticodon and the D-domain sequences are essential determinants for plant cytosolic $tRNA^{Val}$ import into mitochondria. Plant J 34:623–633

Dietrich A, Maréchal-Drouard L, Carneiro V, Cosset A, Small I (1996a) A single base change prevents import of cytosolic $tRNA^{Ala}$ into mitochondria in transgenic plants. Plant J 10:913–918

Dietrich A, Small I, Cosset A, Weil JH, Maréchal-Drouard L (1996b) Editing and import: strategies for providing plant mitochondria with a complete set of functional transfer RNAs. Biochimie 78:518–529

Doersen CJ, Guerrier-Takada C, Altman S, Attardi G (1985) Characterization of an RNase P activity from HeLa cell mitochondria. Comparison with the cytosol RNase P activity. J Biol Chem 260:5942–5949

Dong Y, Yoshitomi T, Hu JF, Cui J (2017) Long noncoding RNAs coordinate functions between mitochondria and the nucleus. Epigen Chromatin 10:41

Dorn GW 2nd (2014) LIPCAR: a mitochondrial lnc in the noncoding RNA chain? Circ Res 114:1548–1550

Dörner M, Altmann M, Pääbo S, Mörl M (2001) Evidence for import of a lysyl-tRNA into marsupial mitochondria. Mol Biol Cell 12:2688–2698

Doucet-Beaupré H, Breton S, Chapman EG, Blier PU, Bogan AE, Stewart DT, Hoeh WR (2010) Mitochondrial phylogenomics of the *Bivalvia* (*Mollusca*): searching for the origin and mitogenomic correlates of doubly uniparental inheritance of mtDNA. BMC Evol Biol 10:50

Duchêne AM, Maréchal-Drouard L (2001) The chloroplast-derived *trnW* and *trnM-e* genes are not expressed in *Arabidopsis* mitochondria. Biochem Biophys Res Commun 285:1213–1216

Duchêne AM, Pujol C, Maréchal-Drouard L (2009) Import of tRNAs and aminoacyl-tRNA synthetases into mitochondria. Curr Genet 55:1–18

Englmeier R, Pfeffer S, Forster F (2017) Structure of the human mitochondrial ribosome studied *in situ* by cryoelectron tomography. Structure 25:1574–1581

Entelis NS, Krasheninnikov IA, Martin RP, Tarassov IA (1996) Mitochondrial import of a yeast cytoplasmic $tRNA^{Lys}$: possible roles of aminoacylation and modified nucleosides in subcellular partitioning. FEBS Lett 384:38–42

Entelis NS, Kieffer S, Kolesnikova OA, Martin RP, Tarassov IA (1998) Structural requirements of $tRNA^{Lys}$ for its import into yeast mitochondria. Proc Natl Acad Sci U S A 95:2838–2843

Entelis NS, Kolesnikova OA, Dogan S, Martin RP, Tarassov IA (2001) 5 S rRNA and tRNA import into human mitochondria. Comparison of *in vitro* requirements. J Biol Chem 276:45642–45653

Entelis N, Brandina I, Kamenski P, Krasheninnikov IA, Martin RP, Tarassov I (2006) A glycolytic enzyme, enolase, is recruited as a cofactor of tRNA targeting toward mitochondria in *Saccharomyces cerevisiae*. Genes Dev 20:1609–1620

Esakova O, Krasilnikov AS (2010) Of proteins and RNA: the RNase P/MRP family. RNA 16:1725–1747

Esseiva AC, Naguleswaran A, Hemphill A, Schneider A (2004) Mitochondrial tRNA import in *Toxoplasma gondii*. J Biol Chem 279:42363–42368

Fang Y, Wu H, Zhang T, Yang M, Yin Y, Pan L, Yu X, Zhang X, Hu S, Al-Mssallem IS, Yu J (2012) A complete sequence and transcriptomic analyses of date palm (*Phoenix dactylifera* L.) mitochondrial genome. PLoS One 7:e37164

Forster AC, Altman S (1990) Similar cage-shaped structures for the RNA components of all ribonuclease P and ribonuclease MRP enzymes. Cell 62:407–409

Frechin M, Senger B, Braye M, Kern D, Martin RP, Becker HD (2009) Yeast mitochondrial Gln-$tRNA^{Gln}$ is generated by a GatFAB-mediated transamidation pathway involving Arc1p-controlled subcellular sorting of cytosolic GluRS. Genes Dev 23:1119–1130

French NS, Norton JD (1997) Structure and functional properties of mouse VL30 retrotransposons. Biochim Biophys Acta 1352:33–47

Furukawa R, Yamada Y, Kawamura E, Harashima H (2015) Mitochondrial delivery of antisense RNA by MITO-porter results in mitochondrial RNA knockdown, and has a functional impact on mitochondria. Biomaterials 57:107–115

Fusté JM, Wanrooij S, Jemt E, Granycome CE, Cluett TJ, Shi Y, Atanassova N, Holt IJ, Gustafsson CM, Falkenberg M (2010) Mitochondrial RNA polymerase is needed for activation of the origin of light-strand DNA replication. Mol Cell 37:67–78

Gao Y, Chen ZY, Wang Y, Liu Y, Ma JX, Li YK (2017) Long non-coding RNA ASncmtRNA-2 is upregulated in diabetic kidneys and high glucose-treated mesangial cells. Exp Ther Med 13:581–587

Geiger J, Dalgaard LT (2017) Interplay of mitochondrial metabolism and microRNAs. Cell Mol Life Sci 74:631–646

Giegé R, Sissler M, Florentz C (1998) Universal rules and idiosyncratic features in tRNA identity. Nucleic Acids Res 26:5017–5035

Glover KE, Spencer DF, Gray MW (2001) Identification and structural characterization of nucleus-encoded transfer RNAs imported into wheat mitochondria. J Biol Chem 276:639–648

Gobert A, Gutmann B, Taschner A, Gossringer M, Holzmann J, Hartmann RK, Rossmanith W, Giege P (2010) A single *Arabidopsis* organellar protein has RNase P activity. Nat Struct Mol Biol 17:740–744

Gold HA, Topper JN, Clayton DA, Craft J (1989) The RNA processing enzyme RNase MRP is identical to the Th RNP and related to RNase P. Science 245:1377–1380
Gowher A, Smirnov A, Tarassov I, Entelis N (2013) Induced tRNA import into human mitochondria: implication of a host aminoacyl-tRNA-synthetase. PLoS One 8:e66228
Greber BJ, Bieri P, Leibundgut M, Leitner A, Aebersold R, Boehringer D, Ban N (2015) Ribosome. The complete structure of the 55S mammalian mitochondrial ribosome. Science 348:303–308
Grewe F, Edger PP, Keren I, Sultan L, Pires JC, Ostersetzer-Biran O, Mower JP (2014) Comparative analysis of 11 *Brassicales* mitochondrial genomes and the mitochondrial transcriptome of *Brassica oleracea*. Mitochondrion 19 Pt B:135–143
Grimes BT, Sisay AK, Carroll HD, Cahoon AB (2014) Deep sequencing of the tobacco mitochondrial transcriptome reveals expressed ORFs and numerous editing sites outside coding regions. BMC Genomics 15:31
Gualberto JM, Mileshina D, Wallet C, Niazi AK, Weber-Lotfi F, Dietrich A (2014) The plant mitochondrial genome: dynamics and maintenance. Biochimie 100:107–120
Gustafsson CM, Falkenberg M, Larsson NG (2016) Maintenance and expression of mammalian mitochondrial DNA. Annu Rev Biochem 85:133–160
Gutmann B, Gobert A, Giege P (2012) PRORP proteins support RNase P activity in both organelles and the nucleus in *Arabidopsis*. Genes Dev 26:1022–1027
Hancock K, Hajduk SL (1990) The mitochondrial tRNAs of *Trypanosoma brucei* are nuclear-encoded. J Biol Chem 265:19208–19215
Hernandez-Cid A, Aguirre-Sampieri S, Diaz-Vilchis A, Torres-Larios A (2012) Ribonucleases P/MRP and the expanding ribonucleoprotein world. IUBMB Life 64:521–528
Holec S, Lange H, Kuhn K, Alioua M, Borner T, Gagliardi D (2006) Relaxed transcription in *Arabidopsis* mitochondria is counterbalanced by RNA stability control mediated by polyadenylation and polynucleotide phosphorylase. Mol Cell Biol 26:2869–2876
Holec S, Lange H, Canaday J, Gagliardi D (2008) Coping with cryptic and defective transcripts in plant mitochondria. Biochim Biophys Acta 1779:566–573
Hollingsworth MJ, Martin NC (1986) RNase P activity in the mitochondria of *Saccharomyces cerevisiae* depends on both mitochondrion and nucleus-encoded components. Mol Cell Biol 6:1058–1064
Holzmann J, Frank P, Loffler E, Bennett KL, Gerner C, Rossmanith W (2008) RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. Cell 135:462–474
Huang L, Mollet S, Souquere S, Le Roy F, Ernoult-Lange M, Pierron G, Dautry F, Weil D (2011) Mitochondria associate with P-bodies and modulate microRNA-mediated RNA interference. J Biol Chem 286:24219–24230
Hwang JY, Han GG, Park JY, Kim EM, An CM, Kang JH, Choi YJ, Kim EB (2016) Complete sequence and polymorphisms of female *Ruditapes philippinarum* (*Mollusca*: *Bivalvia*) mitochondria genome. Mitochondrial DNA A DNA Mapp Seq Anal 27:3462–3463
Ibrahim N, Handa H, Cosset A, Koulintchenko M, Konstantinov Y, Lightowlers RN, Dietrich A, Weber-Lotfi F (2011) DNA delivery to mitochondria: sequence specificity and energy enhancement. Pharm Res 28:2871–2882
Islam MS, Studer B, Byrne SL, Farrell JD, Panitz F, Bendixen C, Møller IM, Asp T (2013) The genome and transcriptome of perennial ryegrass mitochondria. BMC Genomics 14:202
Jagannathan R, Thapa D, Nichols CE, Shepherd DL, Stricker JC, Croston TL, Baseler WA, Lewis SE, Martinez I, Hollander JM (2015) Translational regulation of the mitochondrial genome following redistribution of mitochondrial microRNA in the diabetic heart. Circ Cardiovasc Genet 8:785–802
Jarrous N (2017) Roles of RNase P and its subunits. Trends Genet 33:594–603
Jash S, Adhya S (2011) Suppression of reactive oxygen species in cells with multiple mitochondrial DNA deletions by exogenous protein-coding RNAs. Mitochondrion 11:607–614
Jash S, Chowdhury T, Adhya S (2012) Modulation of mitochondrial respiratory capacity by carrier-mediated transfer of RNA *in vivo*. Mitochondrion 12:262–270

Jung KA, Lee S, Kwak MK (2017) NFE2L2/NRF2 activity is linked to mitochondria and AMP-activated protein kinase signaling in cancers through miR-181c/mitochondria-encoded cytochrome c oxidase regulation. Antioxid Redox Signal 27:945–961

Kamarajan BP, Sridhar J, Subramanian S (2012) *In silico* prediction of microRNAs in plant mitochondria. Int J Bioautomat 16:251–262

Kamenski P, Kolesnikova O, Jubenot V, Entelis N, Krasheninnikov IA, Martin RP, Tarassov I (2007) Evidence for an adaptation mechanism of mitochondrial translation *via* tRNA import from the cytosol. Mol Cell 26:625–637

Kamenski P, Smirnova E, Kolesnikova O, Krasheninnikov IA, Martin RP, Entelis N, Tarassov I (2010) tRNA mitochondrial import in yeast: mapping of the import determinants in the carrier protein, the precursor of mitochondrial lysyl-tRNA synthetase. Mitochondrion 10:284–293

Kapushoc ST, Alfonzo JD, Simpson L (2002) Differential localization of nuclear-encoded tRNAs between the cytosol and mitochondrion in *Leishmania tarentolae*. RNA 8:57–68

Karicheva OZ, Kolesnikova OA, Schirtz T, Vysokikh MY, Mager-Heckel AM, Lombes A, Boucheham A, Krasheninnikov IA, Martin RP, Entelis N, Tarassov I (2011) Correction of the consequences of mitochondrial 3243A>G mutation in the MT-TL1 gene causing the MELAS syndrome by tRNA import into mitochondria. Nucleic Acids Res 39:8173–8186

Kim D, Sung YM, Park J, Kim S, Kim J, Park J, Ha H, Bae JY, Kim S, Baek D (2016) General rules for functional microRNA targeting. Nat Genet 48:1517–1526

Kim KM, Noh JH, Abdelmohsen K, Gorospe M (2017) Mitochondrial noncoding RNA transport. BMB Rep 50:164–174

Kiss T, Filipowicz W (1992) Evidence against a mitochondrial location of the 7-2/MRP RNA in mammalian cells. Cell 70:11–16

Kitow J, Derda AA, Beermann J, Kumarswarmy R, Pfanne A, Fendrich J, Lorenzen JM, Xiao K, Bavendiek U, Bauersachs J, Thum T (2016) Mitochondrial long noncoding RNAs as blood based biomarkers for cardiac remodeling in patients with hypertrophic cardiomyopathy. Am J Physiol Heart Circ Physiol 311:H707–H712

Klemm BP, Wu N, Chen Y, Liu X, Kaitany KJ, Howard MJ, Fierke CA (2016) The diversity of ribonuclease P: protein and RNA catalysts with analogous biological functions. Biomol Ther 6: E27

Klimenko ES, Mileikob VA, Morozkinb ES, Laktionovb PP, Konstantinov YM (2011) Study of DNA import and export in potato (*Solanum tuberosum*) mitochondria using quantitative PCR. Biochem (Moscow) Suppl Ser A Membr Cell Biol 5:170–176

Kolesnikova OA, Entelis NS, Mireau H, Fox TD, Martin RP, Tarassov IA (2000) Suppression of mutations in mitochondrial DNA by tRNAs imported from the cytoplasm. Science 289:1931–1933

Kolesnikova O, Entelis N, Kazakova H, Brandina I, Martin RP, Tarassov I (2002) Targeting of tRNA into yeast and human mitochondria: the role of anticodon nucleotides. Mitochondrion 2:95–107

Kolesnikova OA, Entelis NS, Jacquin-Becker C, Goltzene F, Chrzanowska-Lightowlers ZM, Lightowlers RN, Martin RP, Tarassov I (2004) Nuclear DNA-encoded tRNAs targeted into mitochondria can rescue a mitochondrial DNA mutation associated with the MERRF syndrome in cultured human cells. Hum Mol Genet 13:2519–2534

Kolesnikova O, Kazakova H, Comte C, Steinberg S, Kamenski P, Martin RP, Tarassov I, Entelis N (2010) Selection of RNA aptamers imported into yeast and human mitochondria. RNA 16:926–941

Koley S, Adhya S (2013) A voltage-gated pore for translocation of tRNA. Biochem Biophys Res Commun 439:23–29

Konstantinov YM, Dietrich A, Weber-Lotfi F, Ibrahim N, Klimenko ES, Tarasenko VI, Bolotova TA, Koulintchenko MV (2016) DNA import into mitochondria. Biochem Mosc 81:1044–1056

Koulintchenko M, Konstantinov Y, Dietrich A (2003) Plant mitochondria actively import DNA *via* the permeability transition pore complex. EMBO J 22:1245–1254

Koulintchenko M, Temperley RJ, Mason PA, Dietrich A, Lightowlers RN (2006) Natural competence of mammalian mitochondria allows the molecular investigation of mitochondrial gene expression. Hum Mol Genet 15:143–154

Kren BT, Wong PY, Sarver A, Zhang X, Zeng Y, Steer CJ (2009) MicroRNAs identified in highly purified liver-derived mitochondria may play a role in apoptosis. RNA Biol 6:65–72

Kumar R, Maréchal-Drouard L, Akama K, Small I (1996) Striking differences in mitochondrial tRNA import between different plant species. Mol Gen Genet 252:404–411

Kumarswamy R, Bauters C, Volkmann I, Maury F, Fetisch J, Holzmann A, Lemesle G, de Groote P, Pinet F, Thum T (2014) Circulating long noncoding RNA, LIPCAR, predicts survival in patients with heart failure. Circ Res 114:1569–1575

Laforest MJ, Delage L, Maréchal-Drouard L (2005) The T-domain of cytosolic $tRNA^{Val}$, an essential determinant for mitochondrial import. FEBS Lett 579:1072–1078

Landerer E, Villegas J, Burzio VA, Oliveira L, Villota C, Lopez C, Restovic F, Martinez R, Castillo O, Burzio LO (2011) Nuclear localization of the mitochondrial ncRNAs in normal and cancer cells. Cell Oncol (Dordr) 34:297–305

Lechner M, Rossmanith W, Hartmann RK, Tholken C, Gutmann B, Giege P, Gobert A (2015) Distribution of ribonucleoprotein and protein-only RNase P in eukarya. Mol Biol Evol 32:3186–3193

Lesnik C, Golani-Armon A, Arava Y (2015) Localized translation near the mitochondrial outer membrane: an update. RNA Biol 12:801–809

Leucci E, Vendramin R, Spinazzi M, Laurette P, Fiers M, Wouters J, Radaelli E, Eyckerman S, Leonelli C, Vanderheyden K, Rogiers A, Hermans E, Baatsen P, Aerts S, Amant F, Van Aelst S, van den Oord J, de Strooper B, Davidson I, Lafontaine DL, Gevaert K, Vandesompele J, Mestdagh P, Marine JC (2016) Melanoma addiction to the long non-coding RNA SAMMSON. Nature 531:518–522

Leung AK (2015) The whereabouts of microRNA actions: cytoplasm and beyond. Trends Cell Biol 25:601–610

Leygue E (2007) Steroid receptor RNA activator (*SRA1*): unusual bifaceted gene products with suspected relevance to breast cancer. Nucl Recept Signal 5:e006

Li K, Smagula CS, Parsons WJ, Richardson JA, Gonzalez M, Hagler HK, Williams RS (1994) Subcellular partitioning of MRP RNA assessed by ultrastructural and biochemical analysis. J Cell Biol 124:871–882

Li H, Zhang X, Wang F, Zhou L, Yin Z, Fan J, Nie X, Wang P, Fu XD, Chen C, Wang DW (2016) MicroRNA-21 lowers blood pressure in spontaneous hypertensive rats by upregulating mitochondrial translation. Circulation 134:734–751

Lightowlers RN, Chrzanowska-Lightowlers ZM (2008) PPR (pentatricopeptide repeat) proteins in mammals: important aids to mitochondrial gene expression. Biochem J 416:e5–e6

Lithgow T, Schneider A (2010) Evolution of macromolecular import pathways in mitochondria, hydrogenosomes and mitosomes. Philos Trans R Soc Lond Ser B Biol Sci 365:799–817

Liu TT, Zhu D, Chen W, Deng W, He H, He G, Bai B, Qi Y, Chen R, Deng XW (2013) A global identification and analysis of small nucleolar RNAs and possible intermediate-sized non-coding RNAs in *Oryza sativa*. Mol Plant 6:830–846

Lobos-Gonzalez L, Silva V, Araya M, Restovic F, Echenique J, Oliveira-Cruz L, Fitzpatrick C, Briones M, Villegas J, Villota C, Vidaurre S, Borgna V, Socias M, Valenzuela S, Lopez C, Socias T, Varas M, Diaz J, Burzio LO, Burzio VA (2016) Targeting antisense mitochondrial ncRNAs inhibits murine melanoma tumor growth and metastasis through reduction in survival and invasion factors. Oncotarget 7:58331–58350

Long Y, Wang X, Youmans DT, Cech TR (2017) How do lncRNAs regulate transcription? Sci Adv 3:eaao2110

Lu Q, Wierzbicki S, Krasilnikov AS, Schmitt ME (2010) Comparison of mitochondrial and nucleolar RNase MRP reveals identical RNA components with distinct enzymatic activities and protein components. RNA 16:529–537

Magalhaes PJ, Andreu AL, Schon EA (1998) Evidence for the presence of 5S rRNA in mammalian mitochondria. Mol Biol Cell 9:2375–2382

Mahapatra S, Adhya S (1996) Import of RNA into *Leishmania* mitochondria occurs through direct interaction with membrane-bound receptors. J Biol Chem 271:20432–20437

Mahapatra S, Ghosh S, Bera SK, Ghosh T, Das A, Adhya S (1998) The D arm of $tRNA^{Tyr}$ is necessary and sufficient for import into *Leishmania* mitochondria *in vitro*. Nucleic Acids Res 26:2037–2041

Mahata B, Bhattacharyya SN, Mukherjee S, Adhya S (2005) Correction of translational defects in patient-derived mutant mitochondria by complex-mediated import of a cytoplasmic tRNA. J Biol Chem 280:5141–5144

Mahata B, Mukherjee S, Mishra S, Bandyopadhyay A, Adhya S (2006) Functional delivery of a cytosolic tRNA into mutant mitochondria of human cells. Science 314:471–474

Mahato B, Jash S, Adhya S (2011) RNA-mediated restoration of mitochondrial function in cells harboring a Kearns Sayre syndrome mutation. Mitochondrion 11:564–574

Maida Y, Yasukawa M, Furuuchi M, Lassmann T, Possemato R, Okamoto N, Kasim V, Hayashizaki Y, Hahn WC, Masutomi K (2009) An RNA-dependent RNA polymerase formed by TERT and the RMRP RNA. Nature 461:230–235

Makarova JA, Shkurnikov MU, Wicklein D, Lange T, Samatov TR, Turchinovich AA, Tonevitsky AG (2016) Intracellular and extracellular microRNA: an update on localization and biological role. Prog Histochem Cytochem 51:33–49

Maniataki E, Mourelatos Z (2005) Human mitochondrial $tRNA^{Met}$ is exported to the cytoplasm and associates with the Argonaute 2 protein. RNA 11:849–852

Maréchal-Drouard L, Guillemaut P, Cosset A, Arbogast M, Weber F, Weil JH, Dietrich A (1990) Transfer RNAs of potato (*Solanum tuberosum*) mitochondria have different genetic origins. Nucleic Acids Res 18:3689–3696

Marker C, Zemann A, Terhorst T, Kiefmann M, Kastenmayer JP, Green P, Bachellerie JP, Brosius J, Huttenhofer A (2002) Experimental RNomics: identification of 140 candidates for small non-messenger RNAs in the plant *Arabidopsis thaliana*. Curr Biol 12:2002–2013

Martin RP, Schneller JM, Stahl A, Dirheimer G (1979) Import of nuclear deoxyribonucleic acid coded lysine-accepting transfer ribonucleic acid (anticodon C-U-U) into yeast mitochondria. Biochemistry 18:4600–4605

Matsuda D, Dreher TW (2004) The tRNA-like structure of turnip yellow mosaic virus RNA is a 3′-translational enhancer. Virology 321:36–46

Mercer TR, Neph S, Dinger ME, Crawford J, Smith MA, Shearwood AM, Haugen E, Bracken CP, Rackham O, Stamatoyannopoulos JA, Filipovska A, Mattick JS (2011) The human mitochondrial transcriptome. Cell 146:645–658

Michaelis G, Vahrenholz C, Pratje E (1990) Mitochondrial DNA of *Chlamydomonas reinhardtii*: the gene for apocytochrome b and the complete functional map of the 15.8 kb DNA. Mol Gen Genet 223:211–216

Milani L, Ghiselli F, Maurizii MG, Passamonti M (2011) Doubly uniparental inheritance of mitochondria as a model system for studying germ line formation. PLoS One 6:e28194

Mileshina D, Koulintchenko M, Konstantinov Y, Dietrich A (2011) Transfection of plant mitochondria and *in organello* gene integration. Nucleic Acids Res 39:e115

Miller DL, Martin NC (1983) Characterization of the yeast mitochondrial locus necessary for tRNA biosynthesis: DNA sequence analysis and identification of a new transcript. Cell 34:911–917

Mukherjee S, Basu S, Home P, Dhar G, Adhya S (2007) Necessary and sufficient factors for the import of transfer RNA into the kinetoplast mitochondrion. EMBO Rep 8:589–595

Mukherjee S, Mahata B, Mahato B, Adhya S (2008) Targeted mRNA degradation by complex-mediated delivery of antisense RNAs to intracellular human mitochondria. Hum Mol Genet 17:1292–1298

Mukherjee J, Mahato B, Adhya S (2014) Vesicular transport of a ribonucleoprotein to mitochondria. Biol Open 3:1083–1091

Murcha MW, Kubiszewski-Jakubiak S, Teixeira PF, Gügel IL, Kmiec B, Narsai R, Ivanova A, Megel C, Schock A, Kraus S, Berkowitz O, Glaser E, Philippar K, Maréchal-Drouard L, Soll J, Whelan J (2016) Plant-specific preprotein and amino acid transporter proteins are required for tRNA import into mitochondria. Plant Physiol 172:2471–2490

Nagao A, Suzuki T, Katoh T, Sakaguchi Y, Suzuki T (2009) Biogenesis of glutaminyl-mt $tRNA^{Gln}$ in human mitochondria. Proc Natl Acad Sci U S A 106:16209–16214

Nguyen TT, Brenu EW, Staines DR, Marshall-Gradisnik SM (2014) MicroRNAs in the intracellular space, regulation of organelle specific pathways in health and disease. MicroRNA 3:98–107

Noh JH, Kim KM, Abdelmohsen K, Yoon JH, Panda AC, Munk R, Kim J, Curtis J, Moad CA, Wohler CM, Indig FE, de Paula W, Dudekula DB, De S, Piao Y, Yang X, Martindale JL, de Cabo R, Gorospe M (2016) HuR and GRSF1 modulate the nuclear export and mitochondrial localization of the lncRNA RMRP. Genes Dev 30:1224–1239

Noutsopoulos D, Markopoulos G, Vartholomatos G, Kolettas E, Kolaitis N, Tzavaras T (2010) VL30 retrotransposition signals activation of a caspase-independent and p53-dependent death pathway associated with mitochondrial and lysosomal damage. Cell Res 20:553–562

Oda K, Yamato K, Ohta E, Nakamura Y, Takemura M, Nozato N, Akashi K, Ohyama K (1992) Transfer RNA genes in the mitochondrial genome from a liverwort, *Marchantia polymorpha*: the absence of chloroplast-like tRNAs. Nucleic Acids Res 20:3773–3777

Owen GI (2017) Cancer's new Achilles' heel? Oncotarget 8:55774–55775

Paris Z, Rubio MA, Lukes J, Alfonzo JD (2009) Mitochondrial tRNA import in *Trypanosoma brucei* is independent of thiolation and the Rieske protein. RNA 15:1398–1406

Patrushev M, Kasymov V, Patrusheva V, Ushakova T, Gogvadze V, Gaziev A (2004) Mitochondrial permeability transition triggers the release of mtDNA fragments. Cell Mol Life Sci 61:3100–3103

Patrushev M, Kasymov V, Patrusheva V, Ushakova T, Gogvadze V, Gaziev AI (2006) Release of mitochondrial DNA fragments from brain mitochondria of irradiated mice. Mitochondrion 6:43–47

Pfaff J, Meister G (2013) Argonaute and GW182 proteins: an effective alliance in gene silencing. Biochem Soc Trans 41:855–860

Picardi E, Horner DS, Chiara M, Schiavon R, Valle G, Pesole G (2010) Large-scale detection and analysis of RNA editing in grape mtDNA by RNA deep-sequencing. Nucleic Acids Res 38:4755–4767

Pinti MV, Hathaway QA, Hollander JM (2017) Role of microRNA in metabolic shift during heart failure. Am J Physiol Heart Circ Physiol 312:H33–H45

Posakony JW, England JM, Attardi G (1977) Mitochondrial growth and division during the cell cycle in HeLa cells. J Cell Biol 74:468–491

Pozzi A, Plazzi F, Milani L, Ghiselli F, Passamonti M (2017) SmithRNAs: could mitochondria "bend" nuclear regulation? Mol Biol Evol 34:1960–1973

Puranam RS, Attardi G (2001) The RNase P associated with HeLa cell mitochondria contains an essential RNA component identical in sequence to that of the nuclear RNase P. Mol Cell Biol 21:548–561

Rackham O, Shearwood AM, Mercer TR, Davies SM, Mattick JS, Filipovska A (2011) Long noncoding RNAs are generated from the mitochondrial genome and regulated by nuclear-encoded proteins. RNA 17:2085–2093

Rassow J, Dekker PJ, van Wilpe S, Meijer M, Soll J (1999) The preprotein translocase of the mitochondrial inner membrane: function and evolution. J Mol Biol 286:105–120

Richly E, Leister D (2004) NUMTs in sequenced eukaryotic genomes. Mol Biol Evol 21:1081–1084

Rinehart J, Krett B, Rubio MA, Alfonzo JD, Söll D (2005) *Saccharomyces cerevisiae* Imports the cytosolic pathway for Gln-tRNA synthesis into the mitochondrion. Genes Dev 19:583–592

Rippo MR, Olivieri F, Monsurro V, Prattichizzo F, Albertini MC, Procopio AD (2014) MitomiRs in human inflamm-aging: a hypothesis involving miR-181a, miR-34a and miR-146a. Exp Gerontol 56:154–163

Rivas A, Burzio V, Landerer E, Borgna V, Gatica S, Avila R, Lopez C, Villota C, de la Fuente R, Echenique J, Burzio LO, Villegas J (2012) Determination of the differential expression of mitochondrial long non-coding RNAs as a noninvasive diagnosis of bladder cancer. BMC Urol 12:37

Ro S, Ma HY, Park C, Ortogero N, Song R, Hennig GW, Zheng H, Lin YM, Moro L, Hsieh JT, Yan W (2013) The mitochondrial genome encodes abundant small noncoding RNAs. Cell Res 23:759–774

Rorbach J, Gao F, Powell CA, D'Souza A, Lightowlers RN, Minczuk M, Chrzanowska-Lightowlers ZM (2016) Human mitochondrial ribosomes can switch their structural RNA composition. Proc Natl Acad Sci U S A 113:12198–12201

Rossmanith W (2012) Of P and Z: mitochondrial tRNA processing enzymes. Biochim Biophys Acta 1819:1017–1026

Rossmanith W, Karwan RM (1998) Characterization of human mitochondrial RNase P: novel aspects in tRNA processing. Biochem Biophys Res Commun 247:234–241

Rossmanith W, Tullo A, Potuschak T, Karwan R, Sbisa E (1995) Human mitochondrial tRNA processing. J Biol Chem 270:12885–12891

Rubio MAT, Hopper AK (2011) Transfer RNA travels from the cytoplasm to organelles. Wiley Interdiscip Rev RNA 2:802–817

Rubio MA, Rinehart JJ, Krett B, Duvezin-Caubet S, Reichert AS, Söll D, Alfonzo JD (2008) Mammalian mitochondria have the innate ability to import tRNAs by a mechanism distinct from protein import. Proc Natl Acad Sci U S A 105:9186–9191

Rurek M (2016) Participation of non-coding RNAs in plant organelle biogenesis. Acta Biochim Pol 63:653–663

Rusconi CP, Cech TR (1996) The anticodon is the signal sequence for mitochondrial import of glutamine tRNA in *Tetrahymena*. Genes Dev 10:2870–2880

Russo A, Cirulli C, Amoresano A, Pucci P, Pietropaolo C, Russo G (2008) *Cis*-acting sequences and *trans*-acting factors in the localization of mRNA for mitochondrial ribosomal proteins. Biochim Biophys Acta 1779:820–829

Salinas T, Schaeffer C, Maréchal-Drouard L, Duchêne AM (2005) Sequence dependence of $tRNA^{Gly}$ import into tobacco mitochondria. Biochimie 87:863–872

Salinas T, Duchêne AM, Delage L, Nilsson S, Glaser E, Zaepfel M, Maréchal-Drouard L (2006) The voltage-dependent anion channel, a major component of the tRNA import machinery in plant mitochondria. Proc Natl Acad Sci U S A 103:18362–18367

Salinas T, Duchêne AM, Maréchal-Drouard L (2008) Recent advances in tRNA mitochondrial import. Trends Biochem Sci 33:320–329

Salinas T, Duby F, Larosa V, Coosemans N, Bonnefoy N, Motte P, Maréchal-Drouard L, Remacle C (2012) Co-evolution of mitochondrial tRNA import and codon usage determines translational efficiency in the green alga *Chlamydomonas*. PLoS Genet 8:e1002946

Salinas T, El Farouk-Ameqrane S, Ubrig E, Sauter C, Duchêne AM, Maréchal-Drouard L (2014) Molecular basis for the differential interaction of plant mitochondrial VDAC proteins with tRNAs. Nucleic Acids Res 42:9937–9948

Salinas-Giegé T, Giegé R, Giegé P (2015) tRNA biology in mitochondria. Int J Mol Sci 16:4518–4559

Sbicego S, Nabholz CE, Hauser R, Blum B, Schneider A (1998) *In vivo* import of unspliced $tRNA^{Tyr}$ containing synthetic introns of variable length into mitochondria of *Leishmania tarentolae*. Nucleic Acids Res 26:5251–5255

Schekman R (2010) Editorial expression of concern: a bifunctional tRNA import receptor from *Leishmania* mitochondria. Proc Natl Acad Sci U S A 107:9476

Schneider A (2011) Mitochondrial tRNA import and its consequences for mitochondrial translation. Annu Rev Biochem 80:1033–1053

Schneider A, Maréchal-Drouard L (2000) Mitochondrial tRNA import: are there distinct mechanisms? Trends Cell Biol 10:509–513

Seidman D, Johnson D, Gerbasi V, Golden D, Orlando R, Hajduk S (2012) Mitochondrial membrane complex that contains proteins necessary for tRNA import in *Trypanosoma brucei*. J Biol Chem 287:8892–8903

Sharma A, Sharma A (2015) *Plasmodium falciparum* Mitochondria import tRNAs along with an active phenylalanyl-tRNA synthetase. Biochem J 465:459–469

Shearman JR, Sangsrakru D, Ruang-Areerate P, Sonthirod C, Uthaipaisanwong P, Yoocha T, Poopear S, Theerawattanasuk K, Tragoonrung S, Tangphatsornruang S (2014) Assembly and analysis of a male sterile rubber tree mitochondrial genome reveals DNA rearrangement events and a novel transcript. BMC Plant Biol 14:45

Shepherd DL, Hathaway QA, Pinti MV, Nichols CE, Durr AJ, Sreekumar S, Hughes KM, Stine SM, Martinez I, Hollander JM (2017) Exploring the mitochondrial microRNA import pathway through polynucleotide Phosphorylase (PNPase). J Mol Cell Cardiol 110:15–25

Shi X, Chen DH, Suyama Y (1994) A nuclear tRNA gene cluster in the protozoan *Leishmania tarentolae* and differential distribution of nuclear-encoded tRNAs between the cytosol and mitochondria. Mol Biochem Parasitol 65:23–37

Shinde S, Bhadra U (2015) A complex genome-microRNA interplay in human mitochondria. Biomed Res Int 2015:206382

Sieber F, Duchêne AM, Maréchal-Drouard L (2011) Mitochondrial RNA import from diversity of natural mechanisms to potential applications. Int Rev Cell Mol Biol 287:145–190

Sloan DB, Alverson AJ, Chuckalovcak JP, Wu M, McCauley DE, Palmer JD, Taylor DR (2012) Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. PLoS Biol 10:e1001241

Small I, Maréchal-Drouard L, Masson J, Pelletier G, Cosset A, Weil JH, Dietrich A (1992) *In vivo* import of a normal or mutagenized heterologous transfer RNA into the mitochondria of transgenic plants: towards novel ways of influencing mitochondrial gene expression? EMBO J 11:1291–1296

Smirnov A, Tarassov I, Mager-Heckel AM, Letzelter M, Martin RP, Krasheninnikov IA, Entelis N (2008) Two distinct structural elements of 5S rRNA are needed for its import into human mitochondria. RNA 14:749–759

Smirnov A, Comte C, Mager-Heckel AM, Addis V, Krasheninnikov IA, Martin RP, Entelis N, Tarassov I (2010) Mitochondrial enzyme rhodanese is essential for 5 S ribosomal RNA import into human mitochondria. J Biol Chem 285:30792–30803

Smirnov A, Entelis N, Martin RP, Tarassov I (2011) Biological significance of 5S rRNA import into human mitochondria: role of ribosomal protein MRP-L18. Genes Dev 25:1289–1305

Song S, Jiang F, Yuan J, Guo W, Miao Y (2013) Exceptionally high cumulative percentage of NUMTs originating from linear mitochondrial DNA molecules in the *Hydra magnipapillata* genome. BMC Genomics 14:447

Srinivasan H, Das S (2015) Mitochondrial miRNA (MitomiR): a new player in cardiovascular health. Can J Physiol Pharmacol 93:855–861

Sripada L, Tomar D, Prajapati P, Singh R, Singh AK, Singh R (2012a) Systematic analysis of small RNAs associated with human mitochondria by deep sequencing: detailed analysis of mitochondrial associated miRNA. PLoS One 7:e44873

Sripada L, Tomar D, Singh R (2012b) Mitochondria: one of the destinations of miRNAs. Mitochondrion 12:593–599

Sripada L, Singh K, Lipatova AV, Singh A, Prajapati P, Tomar D, Bhatelia K, Roy M, Singh R, Godbole MM, Chumakov PM, Singh R (2017) Hsa-miR-4485 regulates mitochondrial functions and inhibits the tumorigenicity of breast cancer cells. J Mol Med (Berl) 95:641–651

Stalder L, Heusermann W, Sokol L, Trojer D, Wirz J, Hean J, Fritzsche A, Aeschimann F, Pfanzagl V, Basselet P, Weiler J, Hintersteiner M, Morrissey DV, Meisner-Kober NC (2013) The rough endoplasmatic reticulum is a central nucleation site of siRNA-mediated RNA silencing. EMBO J 32:1115–1127

Stohl LL, Clayton DA (1992) *Saccharomyces cerevisiae* Contains an RNase MRP that cleaves at a conserved mitochondrial RNA sequence implicated in replication priming. Mol Cell Biol 12:2561–2569

Sultan LD, Mileshina D, Grewe F, Rolle K, Abudraham S, Glodowicz P, Niazi AK, Keren I, Shevtsov S, Klipcan L, Barciszewski J, Mower JP, Dietrich A, Ostersetzer-Biran O (2016) The reverse transcriptase/RNA maturase protein MatR is required for the splicing of various group II introns in *Brassicaceae* mitochondria. Plant Cell 28:2805–2829

Sun T, Bentolila S, Hanson MR (2016) The unexpected diversity of plant organelle RNA editosomes. Trends Plant Sci 21:962–973

Suyama Y (1967) The origins of mitochondrial ribonucleic acids in *Tetrahymena pyriformis*. Biochemistry 6:2829–2839

Suyama Y (1986) Two dimensional polyacrylamide gel electrophoresis analysis of *Tetrahymena* mitochondrial tRNA. Curr Genet 10:411–420

Takenaka M, Zehrmann A, Verbitskiy D, Hartel B, Brennicke A (2013) RNA editing in plants and its evolution. Annu Rev Genet 47:335–352

Taliaferro JM, Wang ET, Burge CB (2014) Genomic analysis of RNA localization. RNA Biol 11:1040–1050

Tan T, Pach R, Crausaz Esseiva A, Ivens A, Schneider A (2002a) tRNAs in *Trypanosoma brucei*: genomic organization, expression, and mitochondrial import. Mol Cell Biol 22:3707–3717

Tan TH, Bochud-Allemann N, Horn EK, Schneider A (2002b) Eukaryotic-type elongator $tRNA^{Met}$ of *Trypanosoma brucei* becomes formylated after import into mitochondria. Proc Natl Acad Sci U S A 99:1152–1157

Tarassov I, Entelis N (1992) Mitochondrially-imported cytoplasmic $tRNA^{Lys}$(CUU) of *Saccharomyces cerevisiae*: *in vivo* and *in vitro* targeting systems. Nucleic Acids Res 20:1277–1281

Tarassov I, Entelis N, Martin RP (1995a) An intact protein translocating machinery is required for mitochondrial import of a yeast cytoplasmic tRNA. J Mol Biol 245:315–323

Tarassov I, Entelis N, Martin RP (1995b) Mitochondrial import of a cytoplasmic lysine-tRNA in yeast is mediated by cooperation of cytoplasmic and mitochondrial lysyl-tRNA synthetases. EMBO J 14:3461–3471

Tarassov I, Kamenski P, Kolesnikova O, Karicheva O, Martin RP, Krasheninnikov IA, Entelis N (2007) Import of nuclear DNA-encoded RNAs into mitochondria and mitochondrial translation. Cell Cycle 6:2473–2477

Tonin Y, Heckel AM, Vysokikh M, Dovydenko I, Meschaninova M, Rotig A, Munnich A, Venyaminova A, Tarassov I, Entelis N (2014) Modeling of antigenomic therapy of mitochondrial diseases by mitochondrially addressed RNA targeting a pathogenic point mutation in mitochondrial DNA. J Biol Chem 289:13323–13334

Topper JN, Clayton DA (1990a) Characterization of human MRP/Th RNA and its nuclear gene: full length MRP/Th RNA is an active endoribonuclease when assembled as an RNP. Nucleic Acids Res 18:793–799

Topper JN, Clayton DA (1990b) Secondary structure of the RNA component of a nuclear/mitochondrial ribonucleoprotein. J Biol Chem 265:13254–13262

Tschopp F, Charriere F, Schneider A (2011) *In vivo* study in *Trypanosoma brucei* links mitochondrial transfer RNA import to mitochondrial protein import. EMBO Rep 12:825–832

Val R, Wyszko E, Valentin C, Szymanski M, Cosset A, Alioua M, Dreher TW, Barciszewski J, Dietrich A (2011) Organelle trafficking of chimeric ribozymes and genetic manipulation of mitochondria. Nucleic Acids Res 39:9262–9274

Valach M, Burger G, Gray MW, Lang BF (2014) Widespread occurrence of organelle genome-encoded 5S rRNAs including permuted molecules. Nucleic Acids Res 42:13764–13777

Valentino A, Calarco A, Di Salle A, Finicelli M, Crispi S, Calogero RA, Riccardo F, Sciarra A, Gentilucci A, Galderisi U, Margarucci S, Peluso G (2017) Deregulation of microRNAs mediated control of carnitine cycle in prostate cancer: molecular basis and pathophysiological consequences. Oncogene 36:6030–6040

Vargas JNS, Kar AN, Kowalak JA, Gale JR, Aschrafi A, Chen CY, Gioio AE, Kaplan BB (2016) Axonal localization and mitochondrial association of precursor microRNA 338. Cell Mol Life Sci 73:4327–4340

Vendramin R, Marine JC, Leucci E (2017) Non-coding RNAs: the dark side of nuclear-mitochondrial communication. EMBO J 36:1123–1133

Vidaurre S, Fitzpatrick C, Burzio VA, Briones M, Villota C, Villegas J, Echenique J, Oliveira-Cruz L, Araya M, Borgna V, Socias T, Lopez C, Avila R, Burzio LO (2014) Down-regulation of the antisense mitochondrial non-coding RNAs (ncRNAs) is a unique vulnerability of cancer cells and a potential target for cancer therapy. J Biol Chem 289:27182–27198

Viereck J, Thum T (2017) Circulating noncoding RNAs as biomarkers of cardiovascular disease and injury. Circ Res 120:381–399

Villegas J, Zarraga AM, Muller I, Montecinos L, Werner E, Brito M, Meneses AM, Burzio LO (2000) A novel chimeric mitochondrial RNA localized in the nucleus of mouse sperm. DNA Cell Biol 19:579–588

Villegas J, Araya P, Bustos-Obregon E, Burzio LO (2002a) Localization of the 16S mitochondrial rRNA in the nucleus of mammalian spermatogenic cells. Mol Hum Reprod 8:977–983

Villegas J, Muller I, Arredondo J, Pinto R, Burzio LO (2002b) A putative RNA editing from U to C in a mouse mitochondrial transcript. Nucleic Acids Res 30:1895–1901

Villegas J, Burzio V, Villota C, Landerer E, Martinez R, Santander M, Pinto R, Vera MI, Boccardo E, Villa LL, Burzio LO (2007) Expression of a novel non-coding mitochondrial RNA in human proliferating cells. Nucleic Acids Res 35:7336–7347

Villota C, Campos A, Vidaurre S, Oliveira-Cruz L, Boccardo E, Burzio VA, Varas M, Villegas J, Villa LL, Valenzuela PD, Socias M, Roberts S, Burzio LO (2012) Expression of mitochondrial non-coding RNAs (ncRNAs) is modulated by high risk human papillomavirus (HPV) oncogenes. J Biol Chem 287:21303–21315

Vincent T, Vingadassalon A, Ubrig E, Azeredo K, Srour O, Cognat V, Graindorge S, Salinas T, Marechal-Drouard L, Duchene AM (2017) A genome-scale analysis of mRNAs targeting to plant mitochondria: upstream AUGs in 5′-UTRs reduce mitochondrial association. Plant J 92 (6):1132–1142. https://doi.org/10.1111/tpj.13749

Vinogradova E, Salinas T, Cognat V, Remacle C, Maréchal-Drouard L (2009) Steady-state levels of imported tRNAs in *Chlamydomonas* mitochondria are correlated with both cytosolic and mitochondrial codon usages. Nucleic Acids Res 37:1521–1528

Walker SC, Engelke DR (2008) A protein-only RNase P in human mitochondria. Cell 135:412–414

Wang WX, Springer JE (2015) Role of mitochondria in regulating microRNA activity and its relevance to the central nervous system. Neural Regen Res 10:1026–1028

Wang G, Chen HW, Oktay Y, Zhang J, Allen EL, Smith GM, Fan KC, Hong JS, French SW, McCaffery JM, Lightowlers RN, Morse HC 3rd, Koehler CM, Teitell MA (2010) PNPASE regulates RNA import into mitochondria. Cell 142:456–467

Wang G, Shimada E, Koehler CM, Teitell MA (2012a) PNPASE and RNA trafficking into mitochondria. Biochim Biophys Acta 1819:998–1007

Wang G, Shimada E, Zhang J, Hong JS, Smith GM, Teitell MA, Koehler CM (2012b) Correcting human mitochondrial mutations with targeted RNA import. Proc Natl Acad Sci U S A 109:4840–4845

Wang Y, Wang X, Deng W, Fan X, Liu TT, He G, Chen R, Terzaghi W, Zhu D, Deng XW (2014) Genomic features and regulatory roles of intermediate-sized non-coding RNAs in *Arabidopsis*. Mol Plant 7:514–527

Wang WX, Visavadiya NP, Pandya JD, Nelson PT, Sullivan PG, Springer JE (2015) Mitochondria-associated microRNAs in rat hippocampus following traumatic brain injury. Exp Neurol 265:84–93

Wang J, Meng X, Dobrovolskaya OB, Orlov YL, Chen M (2017) Non-coding RNAs and their roles in stress response in plants. Genomics Proteomics Bioinformatics 15:301–312

Weber-Lotfi F, Koulintchenko MV, Ibrahim N, Hammann P, Mileshina DV, Konstantinov YM, Dietrich A (2015) Nucleic acid import into mitochondria: new insights into the translocation pathways. Biochim Biophys Acta 1853:3165–3181

Weis BL, Schleiff E, Zerges W (2013) Protein targeting to subcellular organelles via mRNA localization. Biochim Biophys Acta 1833:260–273

Woischnik M, Moraes CT (2002) Pattern of organization of human mitochondrial pseudogenes in the nuclear genome. Genome Res 12:885–893

Wu Z, Cuthbert JM, Taylor DR, Sloan DB (2015a) The massive mitochondrial genome of the angiosperm *Silene noctiflora* is evolving by gain or loss of entire chromosomes. Proc Natl Acad Sci U S A 112:10185–10191

Wu Z, Stone JD, Storchova H, Sloan DB (2015b) High transcript abundance, RNA editing, and small RNAs in intergenic regions within the massive mitochondrial genome of the angiosperm *Silene noctiflora*. BMC Genomics 16:938

Yamada Y, Furukawa R, Harashima H (2016) A dual-ligand liposomal system composed of a cell-penetrating peptide and a mitochondrial RNA aptamer synergistically facilitates cellular uptake and mitochondrial targeting. J Pharm Sci 105:1705–1713

Yamamura S, Imai-Sumida M, Tanaka Y, Dahiya R (2017) Interaction and cross-talk between non-coding RNAs. Cell Mol Life Sci 10:1007

Yang W, Wu JM, Bi AD, Ou-Yang YC, Shen HH, Chirn GW, Zhou JH, Weiss E, Holman EP, Liao DJ (2013) Possible formation of mitochondrial-RNA containing chimeric or trimeric RNA implies a post-transcriptional and post-splicing mechanism for RNA fusion. PLoS One 8: e77016

Yang KC, Yamada KA, Patel AY, Topkara VK, George I, Cheema FH, Ewald GA, Mann DL, Nerbonne JM (2014) Deep RNA sequencing reveals dynamic regulation of myocardial non-coding RNAs in failing human heart and remodeling with mechanical circulatory support. Circulation 129:1009–1021

Yoshionari S, Koike T, Yokogawa T, Nishikawa K, Ueda T, Miura K, Watanabe K (1994) Existence of nuclear-encoded 5S-rRNA in bovine mitochondria. FEBS Lett 338:137–142

Zelenka J, Jezek P (2016) Import of fluorescent RNA into mitochondria of living cells. Methods Mol Biol 1351:175–181

Zelenka J, Alan L, Jaburek M, Jezek P (2014) Import of desired nucleic acid sequences using addressing motif of mitochondrial ribosomal 5S-rRNA for fluorescent *in vivo* hybridization of mitochondrial DNA and RNA. J Bioenerg Biomembr 46:147–156

Zhang X, Zuo X, Yang B, Li Z, Xue Y, Zhou Y, Huang J, Zhao X, Zhou J, Yan Y, Zhang H, Guo P, Sun H, Guo L, Zhang Y, Fu XD (2014) MicroRNA directly enhances mitochondrial translation during muscle differentiation. Cell 158:607–619

Zhang Z, Gao W, Long QQ, Zhang J, Li YF, Liu DC, Yan JJ, Yang ZJ, Wang LS (2017) Increased plasma levels of lncRNA H19 and LIPCAR are associated with increased risk of coronary artery disease in a Chinese population. Sci Rep 7:7491

Chapter 5
Control Mechanisms of the Holo-Editosome in Trypanosomes

Jorge Cruz-Reyes, Blaine H. M. Mooers, Vikas Kumar, Pawan K. Doharey, Joshua Meehan, and Luenn Chaparro

Abstract RNA metabolism in the single mitochondrion of trypanosomes and related kinetoplastid protozoa exhibits a unique posttranscriptional maturation of mRNAs by specific U insertion/deletion RNA editing that creates protein-coding sequences in 12 mRNA targets. In *T. brucei*, the editing apparatus includes over 40 proteins and hundreds of small noncoding guide RNAs (gRNAs). The editing machinery faces several challenges besides the need to coordinate its numerous components. These challenges include specific targeting of over 3000 sites in mRNA-gRNA hybrids, faithful discrimination of a large pool of pre-edited, partially edited intermediates and fully edited transcripts in the mitochondrial milieu, and differential control of editing in insect and mammal hosts. However, the basic mechanistic steps that control substrate loading, initiation, and progression of editing are not understood. A growing understanding of the holo-editosome organization offers important clues. The editing holoenzyme is a dynamic aggregate of multi-protein subcomplexes: the "RNA-free" editing enzyme termed RECC and auxiliary RNPs. One RNP is the REH2C subcomplex that includes an RNA helicase. Another subcomplex is RESC that includes two proposed modules: GRBC and REMC. The current model of RNA editing apparatus involves multi-RNP complexes serving as scaffolds that bring together mRNA, gRNA, and the RECC enzyme. Such molecular scaffolds may provide a context for specific mRNA-gRNA annealing, specific site recognition, and editing fidelity and progression. Here, we review protein components in RECC that exhibit differential effects during the life cycle of trypanosomes and specific components of the auxiliary RNPs that may participate in editing control. Notably, variants of RECC and the accessory RNPs have been identified. These findings lead us to propose an updated model of RNA editing, whereby isoforms of enzymatic and nonenzymatic subcomplexes establish "dynamic"

J. Cruz-Reyes (✉) · V. Kumar · P. K. Doharey · J. Meehan · L. Chaparro
Department of Biochemistry and Biophysics, Texas A&M University, College Station, TX, USA
e-mail: cruzrey@tamu.edu

B. H. Mooers
Department of Biochemistry and Molecular Biology, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

J. Cruz-Reyes, M. Gray (eds.), *RNA Metabolism in Mitochondria*, Nucleic Acids and Molecular Biology 34, https://doi.org/10.1007/978-3-319-78190-7_5

functionally distinct holo-editosomes. This should expand the flexibility and specificity of the control mechanisms in RNA editing.

5.1 Introduction

Kinetoplastid protozoa, including trypanosomes, are early-branching eukaryotes with unprecedented mechanisms of gene expression. One such mechanism is a posttranscriptional remodeling of the mitochondrial mRNA transcriptome by the specific insertion and deletion of uridylates. This unique process, discovered in 1986 and coined "RNA editing" by Rob Benne, was readily recognized as a new paradigm in RNA metabolism (Benne et al. 1986). Currently, the term RNA editing is used broadly to represent a large number of posttranscriptional mechanisms, excluding RNA splicing, 5′ capping, and 3′ tail biogenesis, that alter the sequence of the primary transcripts (see in this volume Chaps. 6, 7 and 8). Pioneer studies by the Stuart, Simpson, Sollner-Webb, and Hajduk labs in the 1990s showed that kinetoplastid RNA editing is protein-catalyzed and directed by small noncoding guide RNAs (gRNAs) that exhibit complementarity to fully edited mRNA via canonical and G∘U base pairs (Blum and Simpson 1990; Seiwert et al. 1996; Rusché et al. 1997; Sabatini et al. 1998). In *Trypanosoma brucei*, the etiologic agent of African trypanosomiasis or sleeping sickness, the mRNA transcripts are remodeled at thousands of sites, in reactions requiring hundreds of different gRNA types (Koslowsky et al. 2013). This massive enterprise involves over 40 proteins, most of which can be assigned to a few macromolecular subcomplexes: the ~20S catalytic RNA editing core complex (RECC, also termed the 20S editosome) and auxiliary editing RNPs (Rusché et al. 1997; Stuart et al. 2005; Panigrahi et al. 2008; Weng et al. 2008; Kumar et al. 2016; Tables 5.1 and 5.2, respectively). A growing number of components in the RNA editing apparatus were recently reviewed (Read et al. 2015; Aphasizheva and Aphasizhev 2016; Cruz-Reyes et al. 2016). An outstanding central question in the field is how the extensive editing process is controlled in trypanosomes. The current chapter addresses this question. Different levels of control are expected during the highly complex editing process. The parasite undergoes stage-specific adaptations in the bloodstream of the mammalian host and in the insect transmission vector, the tsetse fly. At each stage, specific or preferential sets of mRNA transcripts are edited. Evidently, strict control mechanisms in mitochondria are needed to prevent "off-targets" including abundant tRNA and rRNA species. Bloodstream-form (BF) and insect-infecting procyclic-form (PF) trypanosomes grow optimally at different temperatures. This could impose different constraints on RNA structure that may influence the targeting by gRNAs and RNA-protein interactions during editing. Recent studies of the RECC enzyme and auxiliary RNPs suggest additional mechanisms of editing control. These include the identification of three variants of RECC with differing site-specificity (Carnes et al. 2011) (Fig. 5.1). Surprisingly, the RECC enzyme, which provides the catalytic center of the holo-editosomes, includes subunits with differential roles in BF and PF

Table 5.1 Proteins of the RECC subcomplex variants

Name	Synonyms		*T. brucei* Gene ID
RECC			
KREPA1	A1	MP81	Tb927.2.2470
KREPA2	A2	MP63	Tb927.10.8210
KREPA3	A2	MP42	Tb927.8.620
KREPA4	A4	MP24	Tb927.10.5110
KREPA5	A5	MP19	Tb927.8.680
KREPA6	A6	MP18	Tb927.10.5120
KREPB4	B4	MP46	Tb927.11.2990
KREPB5	B5	MP44	Tb927.11.940
KREPB6	B6	MP49	Tb927.3.3990
KREPB7	B7	MP47	Tb927.9.5630
KREPB8	B8	MP41	Tb927.8.5690
KREPB9	B9		Tb927.9.4440
KREPB10	B10		Tb927.8.5700
KREN1	N1	REN1	Tb927.1.1690
KREN2	N2	REN2	Tb927.10.5440
KREN3	N3	REN3	Tb927.10.5320
KRET2	T2	RET2	Tb927.7.1550
KREX1	X1	REX1	Tb927.7.1070
KREX2	X2	REX2	Tb927.10.3570
KREL1	L1	REL1	Tb927.9.4360
KREL2	L2	REL2	Tb927.1.3030

trypanosomes (McDermott et al. 2015b). Also, an accessory RNP carries an enzymatic RNA helicase, and nonenzymatic RNP variants have been proposed (Madina et al. 2014; Kumar et al. 2016; Cruz-Reyes et al. 2016; Kafková et al. 2012) (Fig. 5.2). Thus, both core and accessory components of holo-editosomes may contribute to the control of this amazing process. Overall, the current view of RNA editing control seems complex but also brimming with exciting possibilities. Collectively, recent observations from various laboratories discussed below open paths to improve our understanding of novel control mechanisms in trypanosome RNA editing and mitochondrial RNA metabolism in general.

5.2 Variants of the RECC Editing Enzyme: RECC1, RECC2, and RECC3

The initiation step in processes such as those catalyzed by replisomes, ribosomes, spliceosomes, and transcription complexes is typically subject to strict control. Trypanosome RNA editing begins with the recognition of a suitable mRNA/gRNA bimolecular substrate and the ensuing endonucleolytic cleavage at the first

Table 5.2 Protein components of the accessory GRBC, REMC, and REH2C subcomplexes

Name	Function	Motifs	Synonym	*T. brucei* Gene ID
GRBC				
GRBC1	gRNA binding, gRNA stability		GAP2	Tb927.7.2570
GRBC2	gRNA binding, gRNA stability		GAP1	Tb927.2.3800
GRBC3			MRB8620	Tb927.11.16860
GRBC4			MRB5390	Tb11.02.5390
GRBC5		Pentein	MRB11870	Tb927.10.11870
GRBC6			MRB3010	Tb927.5.3010
GRBC7			MRB0880	Tb927.11.9140
REMC				
REMC1			MRB10130	Tb927.10.10130
REMC2			MRB1860	Tb927.2.1860
REMC3			MRB800	Tb927.7.800
REMC4			MRB8180	Tb927.8.8180
REMC5[a]	RNA binding. Processivity of editing		MRB4160	Tb927.4.4160
REMC5A[a]	RNA binding. Processivity of editing		MRB8170	Tb927.8.8170
RGG2	RNA binding. Processivity of editing	RRM, RGG	TbRGG2	Tb927.10.10830
		PhyH	MRB7260	Tb927.9.7260
REH2C				
REH2	RNA helicase, RNA binding	DEAH/RHA, HA2, OB, dsRBD		Tb927.4.1500
^{H2}F1	REH2 adaptor, REH2 stability	C2H2 zinc fingers		Tb927.6.1680
^{H2}F2	RNA binding			Tb927.6.2140
Other characterized proteins not assigned to the complexes above				
REH1	RNA helicase, gRNA exchange	DEAD-box	Mhel61	Tb927.11.8870
RGG1		RGG	TbRGG1	Tb927.3.1820

[a]Paralogs with 77.3% translated sequence identity

(3′ most) editing site on the mRNA strand. Subsequent steps of U-specific addition or removal and RNA ligase-mediated resealing of the cleaved strand complete a full editing cycle at that site (Fig. 5.1c). This basic three-step catalytic cycle repeats at individual sites as editing progresses in a general 3′ to 5′ direction (Seiwert et al. 1996; Cruz-Reyes and Sollner-Webb 1996; Carnes et al. 2017). Together, substrate recognition and mRNA cleavage may mark the physical and functional engagement of the editing apparatus. Therefore, these events may be key checkpoints in the control of RNA editing.

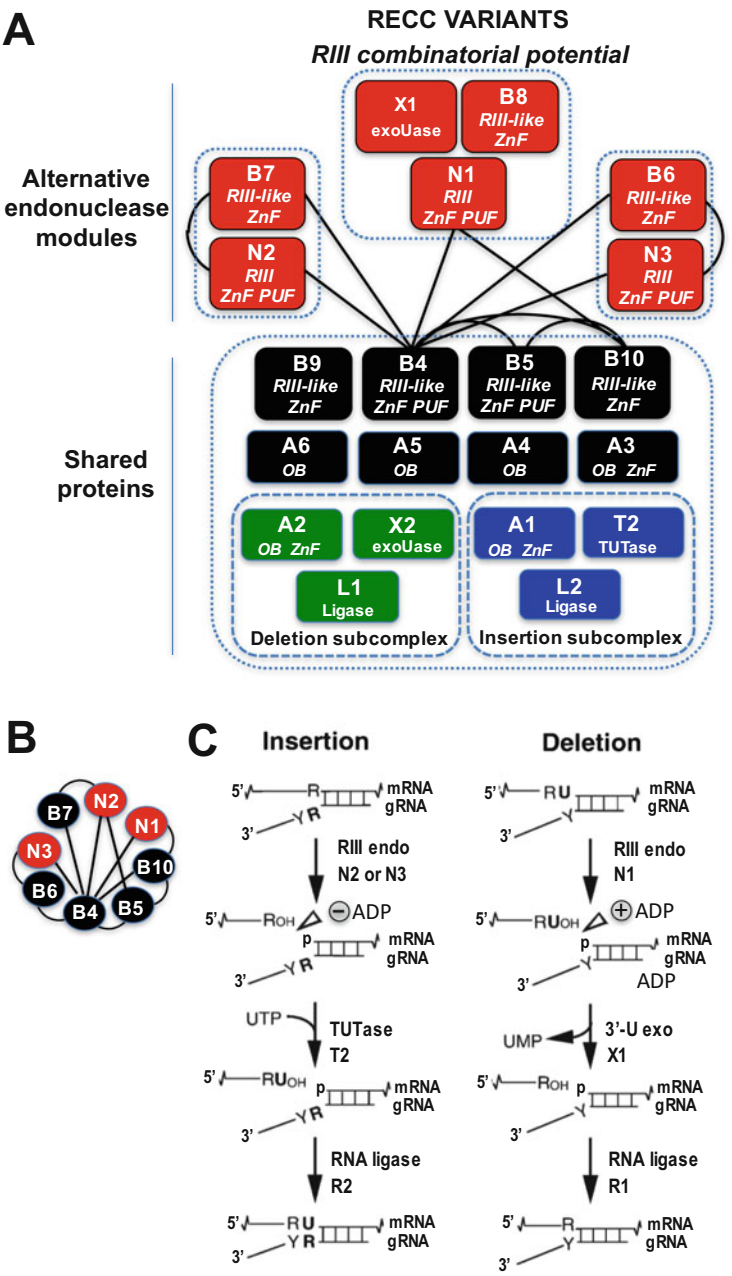

Fig. 5.1 RNA editing core complex (RECC). (**a**) Variants of RECC with shared proteins and alternative endonuclease modules for insertion and deletion editing. BS3 cross-links between RNase III (RIII)-like proteins identified by McDermott et al. (2016) are indicated with lines. (**b**) Summary of the combinatorial potential of RIII-like proteins based on the BS3 cross-linking data in panel (**a**). (**c**) Basic reaction steps at each editing site include gRNA-directed cleavage of the

The first reported purification of the RECC enzyme identified seven protein subunits (Rusché et al. 1997), but subsequent studies nearly tripled this number (Stuart et al. 2005) (Table 5.1). Three of these proteins, REN1 (N1), REN2 (N2), and REN3 (N3), have a functional RNase III-type endonuclease domain. A key for current protein terminology is shown in Table 5.1. Typical RNase III-type nucleases form a homodimer in which both monomers contribute to the creation of a dsRNA-binding surface. The dimer makes a double-stranded RNA cleavage with each monomer cleaving one of the two strands in the RNA duplex (Gan et al. 2008). In contrast, purifications of RECC enzyme exclusively cleave the mRNA strand in mRNA-gRNA duplexes (Seiwert et al. 1996; Rusché et al. 1997; Hernandez et al. 2008). Also, a recombinant version of REN1 cleaved the mRNA strand in an in vitro assay (Kang et al. 2006). Early in vitro studies of gRNA-directed cleavage using mitochondrial extracts and native purified RECC enzyme identified basic requirements for the endonuclease reaction, including a robust stimulation by adenosine nucleotides (ADP or ATP) at deletion sites and their converse inhibition at insertion sites (Cruz-Reyes et al. 1998a, b) (Fig. 5.1c). This early observation revealed a presumed allosteric control of the editing endonucleases and represents the earliest indication of differential control of the cleavage step at sites for insertion and deletion. Nucleotides immediately adjacent to a scissile phosphodiester bond can be manipulated to enhance overall efficiency or artificially convert deletion sites to insertion sites and vice versa (Cruz-Reyes et al. 1998a, 2001; Cifuentes-Rojas et al. 2005). A-form dsRNA is required at, or near, the editing site, and a 2′ hydroxyl is essential at the scissile bond for cleavage (Cifuentes-Rojas et al. 2007). Also, the trypanosomal cleavage activity minimally requires a single helical turn of RNA in the anchor region (Cifuentes-Rojas et al. 2007; Hernandez et al. 2008). While basic requirements for efficient endonucleolytic cleavage in vitro were identified, the discrimination of bona fide editing sites in the mitochondrial milieu leading to a productive cleavage most likely faces additional structural constraints and challenges. The observations above raise central questions, including: how the unique single-stranded cleavage activity of trypanosomal editing nucleases is established, how editing sites are precisely recognized, and how the key editing nuclease step is controlled in vivo.

Genetic studies of the REN proteins in trypanosomes led to important insights regarding these questions. Purification of each tagged endonuclease revealed three variants or isoforms of the RECC enzyme: RECC1, RECC2, and RECC3 (Fig. 5.1a). These variant complexes share a common set of proteins and are distinguished by specific sets of three or two proteins, N1/B8/X1, N2/B7, or N3/B6, respectively (Fig. 5.1a). Reciprocal purifications of tagged B8, B7, and B6 confirmed the

Fig. 5.1 (continued) mRNA, followed by either a 3′-U exonuclease or a 3′ TUTase acting on the cleaved upstream fragment, and RNA ligase action. The gRNA-directed cleavage reactions at U-deletional and U-insertional editing sites are biochemically distinct, with the former requiring an adenosine nucleotide, while the latter is inhibited by adenosine nucleotides

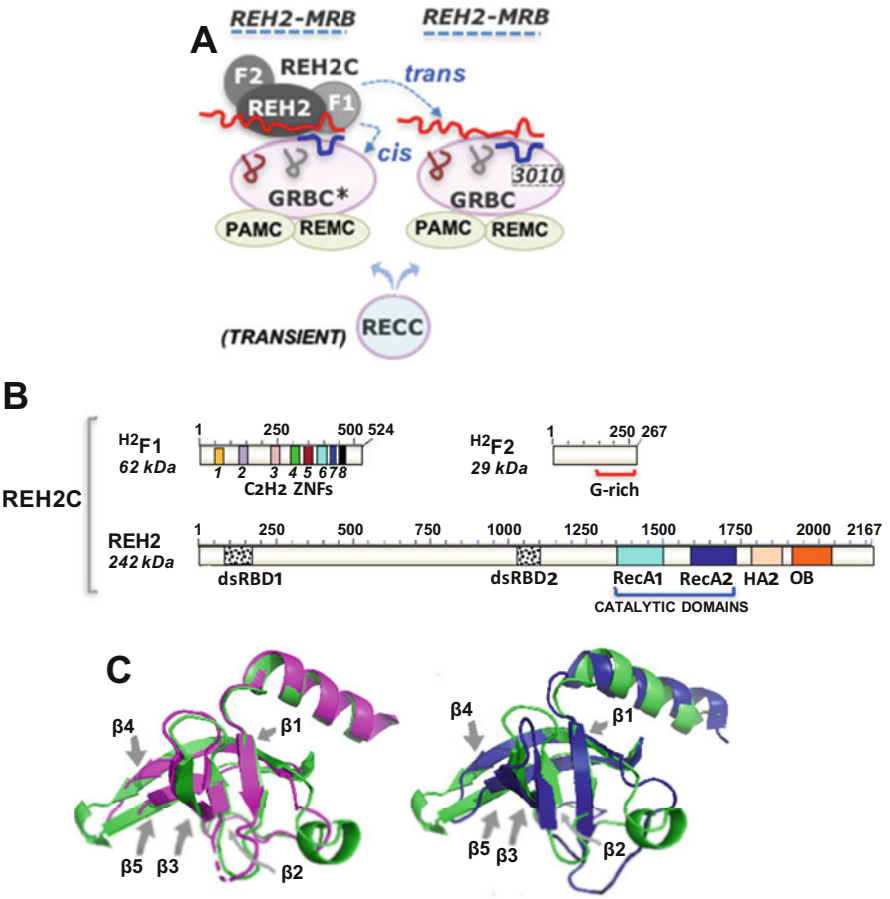

Fig. 5.2 Accessory editing RNPs in holo-editosomes. (**a**) A current model of holo-editosomes whereby accessory multi-RNPs serve as molecular scaffolds that bring together mRNA, gRNA, and the RECC enzyme. The known multi-protein RNPs include the enzymatic REH2C and the nonenzymatic GRBC and REMC. Two GRBC variants discussed here (GRBC and GRBC*) bind and stabilize gRNAs but are readily distinguished by their content of the component MRB3010 ("3010" in the cartoon). The REH2C subcomplex includes the RNA helicase REH2, ^{H2}F1, and ^{H2}F2. The REH2C subcomplex binds stably via RNA with GRBC*/REMC and transiently with GRBC/REMC. Most of the identified proteins were originally thought to form a single complex that was termed MRB1. So, purified complexes using REH2 or MRB3010 (3010) as baits were originally dubbed REH2-MRB and 3010-MRB, respectively. Purifications of GRBC or of REH2C (i.e., in this case from a cell line depleted of cellular gRNA and gRNA-binding GAP1) contain all editing mRNA types: pre-edited, partially edited, and fully edited. The "RNA-free" RECC enzyme (or its variants) transiently interacts with the RNPs via RNA. Components of REMC and of the polyadenylation mediator complex (PAMC) participate in editing progression and post-editing mRNA 3′-end maturation, respectively. The initiating gRNA hybridizes to the first (the 3′ most) block in the pre-edited mRNA. The mRNAs are shown in red, and the gRNA transcripts are shown in various colors. (**b**) Domain organization of the proteins in REH2C. ^{H2}F1 has eight C2H2 zinc fingers. ^{H2}F2 has no bioinformatically identified domains but its C-terminus is glycine-rich. **(c)** (Left) Overlay of OB-fold from chain A of the crystal structure of yeast Prp43 RNA helicase

typifying proteins in each RECC variant (Carnes et al. 2011). Interestingly, B8, B7, and B6 carry one degenerate RNase-like domain that lacks critical catalytic amino acids. Also, the RECC variants exhibit different substrate specificity. Such specificity has been studied in vitro and in vivo (Carnes et al. 2011, 2017). RECC1 primarily cleaves at deletion sites, RECC2 primarily cleaves at insertion sites, and RECC3 primarily cleaves at insertion sites of the mRNA CO2. However, the in vivo studies suggest some flexibility in the specificity of these enzymes at insertion and deletion sites. A current hypothesis is that the single-stranded cleavage activity of the RECC variants is due to heterodimerization of each REN enzyme with a non-catalytic RNase III protein partner. Consistent with this model, recent studies were unable to find evidence for homodimerization of the REN proteins (McDermott et al. 2016). Interestingly, four of the shared proteins in the RECC variants, B4, B5, B9, and B10, also carry degenerate RNase III domains. So, each RECC variant includes one catalytic REN endonuclease and up to five RNase III-like proteins that are catalytically inert. Thus, the possible binary partnerships may include one REN nuclease and one degenerate RNase III protein or two degenerate RNase III-like proteins. Such combinatorial potential of the RNase III protein network may provide fine-tuning while expanding the recognition of editing substrates. The editing machinery must discriminate thousands of editing sites in mitochondria and cleave them efficiently despite the frequent changes in covalent structure of the mRNA and global conformation of the bimolecular mRNA-gRNA structure which changes to adjust to the variable uridine composition of the mRNA as editing progresses (Reifur et al. 2010). The flexibility and ability of the editing machinery to recognize such highly variable substrates could also contribute to the control of differential editing during the life cycle of *T. brucei*.

Recent studies of the RECC variants in procyclic trypanosomes used chemical cross-linking and mass spectrometry to determine proximities between protein subunits in these complexes (McDermott et al. 2015a, 2016). A large network of possible interactions was established using this powerful approach. This review specifically focused on the catalytic and degenerate RNase III proteins. Importantly, REN1, REN2, and REN3 formed no detectable cross-links with each other (Fig. 5.1a, b). This observation is consistent with the exclusive nature of the REN enzymes in the RECC variants. However, inter-cross-links were detected for the typifying pairs N2/B7 and N3/B6, in RECC2 and RECC3, respectively. The RNase III-like protein B4, which is shared in the RECC variants, exhibited the largest number of proximities including with all three REN enzymes and the RNase III-like proteins B5, B6, B7, and B10. The network of detected cross-links between RNase-like proteins is summarized in Fig. 5.1b. The B4 protein showed similar tripartite

Fig. 5.2 (continued) (PDB-ID 2XAU) in magenta on the *T. brucei* homology model in green. Rmsd of 2.53 Å CA atoms of 64 residues. (Right) Overlay of the OB-fold from chain B of the crystal structure of Fly MLE RNA helicase (PDB-ID 5AOR) in dark blue on the *T. brucei* homology model in green. Rmsd of 3.60 Å CA atoms of 72 residues. The five antiparallel β sheets in the OB-fold β barrel are marked with grey arrows

subsets with N2/B7, N3/B6, and N1/B10. In a few cases, no inter-protein cross-links were detected in these studies, for example, between N1 and B8 in the RECC1 module or inter-protein cross-links involving B8 or B9. However, this result may represent a limitation of the BS3 cross-linker used in these studies. This cross-linker has a linker arm of 11.4 Å, which can react with two lysine residues whose alpha carbon atoms are up to 30 Å apart (McDermott et al. 2016). Thus, it is conceivable that bona fide binding surfaces in some proteins lack pairs of lysines separated by appropriate distances and locations for inter-cross-linking. Conversely, proximal cross-links do not necessarily reflect true protein contacts. Valuable insights about the architecture of RECC complexes were generated, but the extent of the combinatorial potential for the editing RNase III proteins would require validation using complementary approaches, including the use of isolated recombinant proteins. In summary, three alternative RECC variants, alongside their potential binary combinations of three catalytic and seven degenerate RNase III proteins, may control key initiating steps in the discrimination and cleavage of editing sites.

5.3 Differential Roles of the RECC Proteins B4, B5, and A3 in Procyclic-Form and Bloodstream-Form Trypanosomes

Genetic studies of B4, B5, and A3 showed that these proteins, which are common to the RECC variants, exhibit differential behaviors in BF and PF cells. As described above, B4 and B5 have degenerate non-catalytic RNase III domains and could potentially dimerize with catalytic or other non-catalytic RNase III editing proteins. A3 also lacks catalytic motifs. However, all three proteins, B4, B5, and A3, carry domains that suggest functional interactions with RNA or proteins (Fig. 5.1a). Initial RNAi-based genetic knockdowns in BF cells showed that B5 and A3 affect the integrity of RECC complexes. Analysis of A3 in PF cells shows a partial effect (Law et al. 2008; Guo et al. 2010; McDermott et al. 2015b). More detailed characterizations using conditional null cell lines also showed that the lack of B5 eliminated the RECC complexes in BF cells but only partially disrupted them in PF cells. Also, the lack of A3 eliminated the RECC complexes in BF cells but only slightly decreased their sedimentation density in PF cells. However, both B5 and A3 are required for RNA editing and cell growth. Furthermore, the substitution of amino acids at selected sites within these proteins differentially affected cell growth, the integrity of the RECC complexes, and RNA editing in the two life stages (McDermott et al. 2015b). Some mutations in various domains of B5 and A3 affected PF cells, BF cells, or both. In a complementary study, a random mutagenesis of the functional domains of B5 identified eight amino acid substitutions that are lethal in BF cells but not in PF cells. Most of these positions were in the degenerate RNase III-like domain, consistent with the idea that this type of domain controls editing (McDermott et al. 2015a). In the case of B4, a recent study showed that this protein

is similar to B5 in that mutations in their degenerate RNase III domain more severely impact the integrity of the RECC complexes in BF than in PF cells. In this study of B4, mutations in the RNase III domain strongly inhibited BF and PF growth and editing (McDermott and Stuart 2017).

Originally, the RECC complexes were thought to largely provide the basic catalytic center in the holo-editosomes. However, the above mutagenic studies provided the first evidence of differential relevance for some subunits of these complexes in BF and PF trypanosomes. So, the RECC complexes may also play a role in the control of stage-specific RNA editing. The above observations also showed that functional studies in one stage are not necessarily valid in another stage. A number of possible mechanisms may account for the differential roles of specific subunits of the RECC complexes including differences in protein modification, conformation, or interactions with other proteins or RNA. Interestingly, the A3 protein is a proposed structural core component of RECC complexes. In PF cells, A3 showed a number of cross-links with specialized components of both the U-insertion and the U-deletion pathways (McDermott et al. 2015a). It will be interesting to compare how the intra-cross-link network of A3, B4, and B5 differ in purified complexes from both PF and BF cells. B4 seems also particularly interesting because it generated in PF cells the largest number of cross-links with other RNase III-type proteins. Yet another RNase III protein, like B5, or another core protein besides A3 (i.e., also shared in the RECC variants), may be a major interaction point in BF cells. It is possible that RECC complexes in BF and PF cells exhibit a similar protein composition. However, the low expression of tagged proteins in BF have prevented isolation of sufficient RECC complexes for analysis in this life cycle stage (Carnes et al. 2011). Even if RECC complexes are compositionally identical in BF and PF cells, the dramatic differences caused by the mutations described above suggest that the architecture of the RECC complexes, or their protein-protein and protein-RNA interactions, differ substantially in the two stages.

5.4 Enzymatic and Nonenzymatic Auxiliary RNPs of the Editing Apparatus: REH2C, GRBC, and REMC

The current assignment of individual proteins to the known editing subcomplexes is mostly based on RNA-independent co-purification, direct physical interaction of isolated proteins, or both. Also, it is common to find that one protein stabilizes another protein(s) in the same subcomplex. Two subcomplexes are discussed here: the RNA editing helicase 2 complex (REH2C) and the RNA editing substrate complex (RESC). Most of the proteins that currently define the REH2C and RESC were initially detected in early purifications of a proposed higher-order RNP originally termed mitochondrial RNA-binding complex 1 (MRB1) or gRNA-binding complex (GRBC) (Panigrahi et al. 2008; Hashimi et al. 2008; Weng et al. 2008; Hernandez et al. 2010). The interdependence for stability among some proteins in

the subcomplexes helped define two proposed modules of RESC: the core GRBC and the more loosely defined RNA editing mediator complex (REMC) (Aphasizheva et al. 2014). Protein components of GRBC and REMC stably co-purify in immunoprecipitations or affinity purifications of accessory editing proteins (Weng et al. 2008; Hernandez et al. 2010; Ammerman et al. 2012). However, as discussed below, GRBC and REMC appear to have different editing functions. The dependence for protein stability in the subcomplexes is not always mutual: that is, one protein may stabilize another protein but not vice versa. One protein subunit may specifically affect the function(s) of another by directly impacting its stability. However, the loss of a core component could disrupt the assembly or the stability of the entire subcomplex. Interestingly, variants of GRBC have been reported that differ in their content of gRNA and one or more proteins (Madina et al. 2015). There is also evidence that variants of REMC exist (Kafková et al. 2012). Putative variants of the editing subcomplexes may exhibit overlapping but differential functions that increase flexibility in function and specificity during the extensive remodeling of the mitochondrial transcriptome. A central concept in understanding the function of enzymatic and nonenzymatic RNPs in editing derives from the realization that these RNPs are enriched of editing mRNAs, both substrates, and products. Such data led to a model in which the auxiliary editing RNPs serve as scaffolds for a dynamic assembly of mRNA-gRNA hybrid substrates and active RECC enzyme (Madina et al. 2014; Aphasizheva et al. 2014).

5.5 The REH2C Helicase Subcomplex and Its Association with Variants of GRBC

The ~15S REH2C was first identified in purifications of the RNA helicase REH2 from RNA-depleted mitochondria after a genetic knockdown of the single T7 bacteriophage-like RNA polymerase in the organelle (Madina et al. 2015). The isolated REH2C includes the REH2 RNA helicase, REH2-associated factor 1 (^{H2}F1), and ^{H2}F2A (Kumar et al. 2016). A purification of REH2C from gRNA-depleted cells after the knockdown of a core protein in GRBC, the gRNA-binding protein GAP1, showed that REH2C retains fully edited mRNA, pre-edited precursors, and partially edited intermediates (Kumar et al. 2016). Thus, REHC2 is an mRNP that includes three protein subunits and all mRNA types that are involved in editing. These editing mRNAs in REH2C are enriched in the subcomplex relative to never-edited mRNAs in mitochondria and relative to nuclear and cytosolic transcripts (Kumar et al. 2016).

In the REH2C subcomplex, only REH2 and ^{H2}F1 are required for editing in procyclic trypanosomes. Interestingly, a knockdown of ^{H2}F1 affects the REH2 in at least two ways in vivo: it hinders REH2 co-purification with RESC components (i.e., in GRBC/REMC), and it causes REH2 fragmentation. The ^{H2}F1 knockdown does not prevent the association of REH2 with ^{H2}F2. Thus, ^{H2}F1 is an adaptor protein that stabilizes REH2 and brings it to the editing apparatus (Kumar et al. 2016). In

agreement with a stabilizing role, the overexpression of ^{H2}F1 has been seen to moderately increase the REH2 level at steady state (Kumar et al. In Preparation). The immunoprecipitated native REH2C exhibits 3′–5′ ATP-dependent unwinding activity (Hernandez et al. 2010). The isolated recombinant REH2 also catalyzes ATP-dependent unwinding of dsRNA in vitro. This observation indicates that REH2 may indeed function as an ATP-dependent RNA helicase in RNA editing (Kumar et al. In preparation). REH2C associates with the gRNA-bound GRBC via RNA contacts. Interestingly, different types of RNA-based association between these two RNPs have been observed. This result led to a model of the holo-editosome including two variants of the GRBC, namely, GRBC and GRBC* (Madina et al. 2014, 2015; Kumar et al. 2016) (Fig. 5.2). Notably, REH2C binds stably with GRBC*, whereas the interaction of REH2C with GRBC is transient. Both types of association by REHC2, stable or transient, are mediated by RNA. GRBC and GRBC* are also readily distinguished by their differential content of the protein subunit MRB3010. That is, relative to the core GAP1/GAP2 proteins that exist in both GRBC variants, MRB3010 is readily detected in GRBC but not in GRBC*. It is unclear whether GRBC* and GRBC have other distinguishing protein marker(s). However, the GRBC variants can be separated via purification of REH2 or ^{H2}F1 (in the case of GRBC*) or via purification of MRB3010 (in the case of GRBC). Both the purified REH2C•GRBC* co-complex (via REH2) and GRBC (via MRB3010) are enriched with fully edited mRNA transcripts compared to total mtRNA (Madina et al. 2014). A separate study also showed that GRBC associates with editing mRNAs (Aphasizheva et al. 2014). Thus, the independent studies by Madina et al. (2014) and Aphasizheva et al. (2014) provided the first experimental data supporting a model of the holo-editosome in which accessory RNPs bring together mRNA, gRNA, and the RECC enzyme. Other studies have compared GRBC and GRBC*, including their RNA content and interactions with other components of the editing machinery. An RNA-Seq study that examined the content of initiating gRNAs found that GRBC exhibits a relative enrichment of several initiating gRNAs (e.g., gRNAs covering the first 1–3 blocks at the 3′ end of an editing domain) compared to the stable REH2C•GRBC* co-complex and total mtRNA (Madina et al. 2014). Interestingly, MRB3010 is thought to participate early in the editing process (Ammerman et al. 2011). Manual sequencing of cDNA clones and qRT-PCR of the first editing block in mRNAs A6 and ND7, or the first few blocks in mRNA RPS12, indicated a greater accuracy in editing by the initiating gRNA in the mRNA found in GRBC than in REH2C•GRBC* or in total mtRNA (Madina et al. 2015). Each gRNA directs the editing of one block (Koslowsky et al. 2013). Thus, the MRB3010-containing GRBC variant appears to support relatively more efficient editing by RECC than the REH2C•GRBC* co-complex. However, both REH2 and ^{H2}F1, in the REH2C subcomplex, cause *trans* effects on GRBC. For example, the REH2 knockdown reduced the editing of mRNAs associated with GRBC. This editing inhibition was observed in qRT-PCR assays of either the initial 3′ block or a distal 5′ block in the mRNA transcripts. This finding suggests that REH2 affects most, if not all, editing sites during early and late stages of editing progression. The REH2 knockdown also induced increased pausing at preferential sites in the mRNA transcripts examined

(Madina et al. 2015). It is possible that REH2 remodels the mRNA-gRNA local structure at each site to improve access by RECC. Similarly, mRNA transcripts that associate with GRBC are also less completely edited upon ^{H2}F1 depletion. In the ^{H2}F1 knockdown, pulldowns of GRBC (via MRB3010) exhibited a reduced level of associated RECC enzyme and pre-edited mRNA (Madina et al. 2015; Kumar et al. 2016; Kumar et al. unpublished data). This result suggests that the REH2C subcomplex can affect GRBC *in trans*. Specifically, the recruitment of REH2 by the adaptor protein ^{H2}F1 may promote binding or retention of pre-edited mRNA and RECC enzyme by GRBC. Moreover, band shift assays showed that immunoprecipitated GRBC and REH2C•GRBC* form a RNP with a short radioactive RNA duplex that is slow-moving in native gels (Kumar et al. 2016). Interestingly, depletion of ^{H2}F1 in mitochondrial extracts reduced the level of this RNP in both isolated subcomplexes. The observations above indicate that REH2C acts *in trans* to influence GRBC in multiple ways. Thus, dynamic physical and functional interactions of REH2C with variants of GRBC, and with variants of RECC (described above), could increase the combinatorial potential of the subcomplex variants in holo-editosomes.

5.6 A New Class of Protein Regulators of DExH/RHA RNA Helicases Includes the Trypanosomal ^{H2}F1 and Other Multi-zinc Finger Proteins in Evolutionarily Distant Eukaryotes

The control of helicase function in DExH/RHA-subfamily members often involves the direct interaction between the characteristic OB-fold in the C-terminus of this type of helicases and factors termed G-path proteins (Robert-Paganin et al. 2015). DExH/RHA helicases from yeast, flies, and humans participate in diverse RNA processes including mRNA splicing, ribosome biogenesis, and X-chromosome activation. Studies of the OB-fold in these helicases indicate that this conserved domain is regulatory (Walbott et al. 2010; Jarmoskaite and Russell 2014; Prabu et al. 2015). However, bioinformatic domain searches failed to identify recognizable G-patch domains in the trypanosomal ^{H2}F1 or ^{H2}F2 (Kumar et al. 2016). No crystal structure of a G-patch is currently available (Robert-Paganin et al. 2015). However, the C-terminal half of REH2 binds directly with ^{H2}F1 (Kumar et al. 2016), and this interaction involves a ~250 amino acid fragment including the OB-fold in REH2 (Doharey et al., unpublished data). We recently proposed that besides typical G-patch proteins, there is a second class of DExH/RHA helicase protein regulators that contain multiple zinc fingers (Cruz-Reyes et al. 2016). Examples of other DExH/RHA helicase•Znf cofactor systems include the following cases in nematodes and humans. In the gonads of hermaphroditic worms, the nuclear MOP•MEP-1 system is thought to control the spliceosome-mediated location of EJCs near exon-exon junctions in target mRNAs involved in the sperm-oocyte switch (Bono and Gehring 2011). In humans, the

cytosolic DXH30•ZAP system specifically binds viral mRNA targets and directs them to exosomes for their specific degradation (Guo et al. 2007; Ye et al. 2010). So, these helicase•Znf protein systems, including the REH2•^{H2}F1 pair, control the function of specialized mRNPs. ^{H2}F1, MEP-1, and ZAP seem to be dimeric (Bono and Gehring 2011; Ye et al. 2010; Kumar et al. Unpublished data). Another common feature of these zinc finger proteins is that they all serve as adaptors directing their helicase partners to editosomes, spliceosomes, and exosomes, respectively. The evolutionary distance between early-branching trypanosomes and vertebrates suggests that zinc finger cofactors likely modulate DExH/RHA helicases in additional RNA processes where they may be present. The high conservation of the OB-fold enables detailed sequence and structure models of this domain in REH2, as we have found using available crystal structures of RNA helicases in yeast and flies, Prp43p and MLE, respectively (Fig. 5.2c). We are currently using this information to probe conserved features that may affect helicase function in REH2 in trypanosomes and evolutionarily distant RNA helicases (Kumar et al. submitted for publication).

5.7 Different Contribution to Editing Progression by Protein Components of REMC

Genetic studies of proteins in the auxiliary RNPs often show differences in the impact of these proteins on the editing efficiency of different mRNA substrates. However, the functional basis of those differences remains unclear in most cases. This is further complicated because most of these proteins lack bioinformatically identifiable motifs. Also, detailed mechanistic studies are necessary. A study of protein components of REMC, namely, RGG2 (aka TbRGG2), REMC4 (MRB8180), and the 77% identical paralogs REMC5A/REMC5 (MRB4170/MRB4160), found that these proteins exhibit differential effects on editing progression in two examined mRNAs (Simpson et al. 2017). The paralogs REMC5A and REMC5 are redundant, so a double knockdown (dKD) of these proteins was necessary to induce a substantial editing phenotype in procyclic trypanosomes. In addition to the distinct effect of individual REMC subunits, it is feasible that variants of REMC occur in vivo. In that case, the function and specificity of this module could be more dynamic. Consistent with this idea, a RGG2 knockdown reduced the levels of REMC5A but had no effect on the steady-state level of REMC4. In contrast, knockdowns of REMC4 or REMC5A/REMC5 had no effect on the level of RGG2 (Simpson et al. 2017). Notably, affinity purifications of tagged REMC5A and REMC5 paralogs showed similar, but not identical, sets of associated proteins. In fact, the tagged purification of either paralog had an underrepresentation of unique peptides attributable to the other paralog. Furthermore, the REMC5A and REMC5 paralogs failed to co-purify in samples treated with RNase, and the knockdown of one paralog did not affect the isopycnic sedimentation of the other (Kafková et al. 2012). So, REMC5A and REMC5 are part of distinct variants of the REMC module.

Together, the occurrence of REMC variants and a distinct contribution of individual protein subunits to editing progression increase the potential for differential control. RGG2, REMC4, and REMC5A/REMC5 are RNA-binding proteins, as established in UV cross-linking assays of each recombinant protein with synthetic RNA transcripts (Fisk et al. 2008; Kafková et al. 2012). Genetic knockdowns in trypanosomes suggest that RGG2, REMC4, and REMC5A/REMC5 specifically affect mRNAs that require extensive editing. That is, these knockdowns did not affect mRNAs that require minimal editing or the steady-state level of pre-edited mRNA precursors or transcripts that are never edited. Interestingly, the studies by Simpson et al. also showed differential effects on the editing progression of two mRNAs examined, RPS12 and the ND7–5′ domain (i.e., the mRNA ND7 has separate 5′ and 3′ editing domains, not a single domain as in other mRNAs). Notably, depletion of RGG2 or REMC4 caused pausing at numerous sites on RPS12, but at only a few sites on the ND7–5′ domain. However, the double REMC5A/REMC5 knockdown resulted in the converse effect, that is, there was pausing at numerous sites in the ND7–5′ domain but only at a few sites in RPS12. Importantly, these pauses are spread out on sequences that are targeted by multiple gRNAs. A broad distribution of pausing sites is expected if the editing progression is affected. RGG2 and REMC4 affect mostly the same sites, but a few sites seem specific to one protein or the other. So, RGG2 and REMC4 may have synergistic effects on editing. In comparison to the studies of editing progression where the pools of gRNA may be normal, the genetic knockdown of a core protein (GAP1) caused abrupt pausing specifically at the end of the first or second gRNA. This result was due to the concurrent loss of the GRBC integrity and gRNA in the cell. The assembly of components that form REMC and other editing complexes is an interesting question because it may be controlled in vivo. However, studies of the assembly process in vivo may be challenging because depletions of individual proteins may induce secondary effects on overall complex integrity. Thus, additional control mechanisms that involve the accessory RNPs are emerging including the likely occurrence of REMC variants and the substrate specificity of the REMC subunits during editing progression. This complexity further increases the potential for differential control in the editing machinery.

5.8 Differences in RNA-Binding Preference by Protein Components in REMC

Most, if not all, editing protein components of the RECC enzyme and the auxiliary RNPs may influence, directly or indirectly, the interactions of the editing apparatus with RNA substrates and products. Such protein-RNA interactions may be structural or regulatory. Either situation may control the editing reactions, RNA stability, or the coordination of editing with transcription and translation in mitochondria. A recent study applied UV-cross-linking and affinity purification (iCLAP) techniques to

investigate direct RNA interactions of the paralogs REMC5A/REMC5 (MRB8170/MRB4160) in procyclic trypanosomes (Dixit et al. 2017). REMC5A and REMC5 interacted with all types of mitochondrial mRNA (not with gRNA) including pre-edited and fully edited mRNA, but also with transcripts that are never edited. The study reported the total number of the iCLAP tags (each tag includes ~30–50 nt) for several transcripts. In general, REMC5A and REMC5 exhibited similar binding profiles in their mRNA preferences and the total number of interactions in each transcript. The nature of the editing process and the short length of the tag sequences prevented a clear distinction of whether a fully edited tag derived from the final edited product or intermediates with partial editing (i.e., with an edited 3′ block). Likewise, a pre-edited tag could derive from the pre-edited primary transcript or intermediates with partial editing (i.e., with a pre-edited 5′ block). However, most tags included an extensively edited sequence (~90%), and the least abundant contacts were with never-edited sequences. The highly diverse pool of mRNA molecules in mitochondria requires additional studies. However, a functional interpretation of the specific RNA-protein contacts is beginning to emerge. Full gene maps of the tags comparing two pre-edited transcripts revealed contacts of REMC5A and REMC5 throughout most of the length of the COX3 mRNA. In contrast, both proteins exhibited negligible contacts with a different (ND3) mRNA. Interestingly, a double knockdown (dKD) of REMC5A and REMC5 caused a decrease in the steady-state level of both pre-edited and edited CO3 mRNA. This result was found in two separate studies by the same lab (Kafková et al. 2012; Dixit et al. 2017). This observation suggests that REMC5A and REMC5 specifically stabilize the CO3 mRNA, both before and after editing. The ability to control the steady-steady level of a pre-mRNA would determine how much substrate enters the editing pathway. In other cases, binding of REMC5A and REMC5 to a mRNA transcript could specifically stimulate the editing of that transcript. This may be the case for the A6 mRNA, where dKD of the paralogs consistently increased the level of pre-edited transcript while decreasing the level of edited molecules (Kafková et al. 2012; Dixit et al. 2017). Thus, RNA binding by REMC5A and REMC5 may have differential substrate-specific roles in RNA stability or RNA editing. The precise distribution of these roles in the mitochondrion is somewhat unclear because one study suggested a specific increase in pre-edited mRNAs without changes in their edited counterparts, whereas another study suggested changes in editing (i.e., gain of substrate and loss of product) (Kafková et al. 2012; Dixit et al. 2017).

5.9 Conclusion

Uncovering the mechanisms that control the RNA editing process in trypanosome mitochondria remains a daunting task. However, a recent model of holo-editosomes as dynamic aggregates of specialized subcomplexes offers promising insights. Central to this model is the realization that accessory subcomplexes serve as molecular scaffolds that bring together mRNA, gRNA, and the RECC enzyme.

Finally, the existence of editing subcomplex variants or isoforms also provides critical insights. Indeed, variants of RECC, GRBC, and REMC have been found. Together, these variant subcomplexes may provide the flexibility and fine-tuning that is necessary for the specific recognition of thousands of editing sites in pre-edited substrates and the myriad of intermediates in the mitochondrial milieu. Because the cleavage step by REN nucleases in RECC may be a key checkpoint in editing control, the cross-linking data suggesting a combinatorial potential of catalytic and non-catalytic RNase III proteins are exciting. The future identification and verification of bona fide partners in the putative RNase III heterodimers may establish alternative pairs between catalytic and non-catalytic RNase III proteins and between non-catalytic RNase III proteins and possible roles of the heterodimers. The isolation of editing complexes from bloodstream-form trypanosomes remains problematic due to the difficulty in culturing this life cycle form at high density. Nonetheless, cross-linking maps of editing complexes and the determinations of the protein composition and organization of editing complexes in BF trypanosomes will be important. The stepwise assembly of the individual subcomplexes and their co-complexes may also be controlled based on the metabolic needs of the growing parasites or their adaptation to different hosts. Many questions remain, including the binding and stabilization of gRNA by core proteins of GRBC, the targeting of specific mRNAs by the GRBC, the possible roles of REH2C in remodeling of mRNA-gRNA hybrids for efficient editing by RECC, the coordination of REMC with other subcomplexes during editing progression, the determinants for RNA binding by components of REMC, and the differential effects in stability or editing of their ligands. While trypanosome RNA editing remains a surprisingly intricate process in mitochondria, the evolving model of the holo-editosome organization is brimming with exciting possibilities.

Acknowledgments This work was supported by the National Science Foundation [Collaborative Research 1616865 to J.C.-R. and 1616845 to B.H.M.M.]; AgriLife at TAMU [to J.C.-R.]; and Presbyterian Health Foundation [to B.H.M.M.].

References

Ammerman ML, Hashimi H, Novotná L, Čičová Z, McEvoy SM, Lukeš J, Read LK (2011) MRB3010 is a core component of the MRB1 complex that facilitates an early step of the kinetoplastid RNA editing process. RNA 17:865–877

Ammerman ML, Downey KM, Hashimi H, Fisk JC, Tomasello DL, Faktorová D, Kafková L, King T, Lukeš J, Read LK (2012) Architecture of the trypanosome RNA editing accessory complex, MRB1. Nucleic Acids Res 40:5637–5650

Aphasizheva I, Aphasizhev R (2016) U-insertion/deletion mRNA-editing holoenzyme: definition in sight. Trends Parasitol 32:144–156

Aphasizheva I, Zhang L, Wang X, Kaake RM, Huang L, Monti S, Aphasizhev R (2014) RNA binding and core complexes constitute the U-insertion/deletion editosome. Mol Cell Biol 34:4329–4342

Benne R, Van den Burg J, Brakenhoff JP, Sloof P, Van Boom JH, Tromp MC (1986) Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. Cell 46:819–826

Blum B, Simpson L (1990) Guide RNAs in kinetoplastid mitochondria have a nonencoded 3′ oligo (U) tail involved in recognition of the preedited region. Cell 62:391–397

Bono F, Gehring NH (2011) Assembly, disassembly and recycling: the dynamics of exon junction complexes. RNA Biol 8:24–29

Carnes J, Soares CZ, Wickham C, Stuart K (2011) Endonuclease associations with three distinct editosomes in *Trypanosoma brucei*. J Biol Chem 286:19320–19330

Carnes J, McDermott S, Anupama A, Oliver BG, Sather DN, Stuart K (2017) *In vivo* cleavage specificity of *Trypanosoma brucei* editosome endonucleases. Nucleic Acids Res 45:4667–4686

Cifuentes-Rojas C, Halbig K, Sacharidou A, De Nova-Ocampo M, Cruz-Reyes J (2005) Minimal pre-mRNA substrates with natural and converted sites for full-round U insertion and U deletion RNA editing in trypanosomes. Nucleic Acids Res 33:6610–6620

Cifuentes-Rojas C, Pavia P, Hernandez A, Osterwisch D, Puerta C, Cruz-Reyes J (2007) Substrate determinants for RNA editing and editing complex interactions at a site for full-round U insertion. J Biol Chem 282:4265–4276

Cruz-Reyes J, Sollner-Webb B (1996) Trypanosome U-deletional RNA editing involves guide RNA-directed endonuclease cleavage, terminal U exonuclease, and RNA ligase activities. Proc Natl Acad Sci U S A 93:8901–8906

Cruz-Reyes J, Rusché LN, Piller KJ, Sollner-Webb B (1998a) *T. brucei* RNA editing: adenosine nucleotides inversely affect U-deletion and U-insertion reactions at mRNA cleavage. Mol Cell 1:401–409

Cruz-Reyes J, Rusché LN, Sollner-Webb B (1998b) *Trypanosoma brucei* U insertion and U deletion activities co-purify with an enzymatic editing complex but are differentially optimized. Nucleic Acids Res 26:3634–3639

Cruz-Reyes J, Zhelonkina A, Rusche L, Sollner-Webb B (2001) Trypanosome RNA editing: simple guide RNA features enhance U deletion 100-fold. Mol Cell Biol 21:884–892

Cruz-Reyes J, Mooers BH, Abu-Adas Z, Kumar V, Gulati S (2016) DEAH-RHA helicase•Znf cofactor systems in kinetoplastid RNA editing and evolutionarily distant RNA processes. RNA Dis 3:e1336

Dixit S, Muller-McNicoll M, David V, Zarnack K, Ule J, Hashimi H, Lukes J (2017) Differential binding of mitochondrial transcripts by MRB8170 and MRB4160 regulates distinct editing fates of mitochondrial mRNA in trypanosomes. MBio 8:e02288-16

Fisk JC, Ammerman ML, Presnyak V, Read LK (2008) TbRGG2, an essential RNA editing accessory factor in two *Trypanosoma brucei* life cycle stages. J Biol Chem 283:23016–23025

Gan J, Shaw G, Tropea JE, Waugh DS, Court DL, Ji X (2008) A stepwise model for double-stranded RNA processing by ribonuclease III. Mol Microbiol 67:143–154

Guo X, Ma J, Sun J, Gao G (2007) The zinc-finger antiviral protein recruits the RNA processing exosome to degrade the target mRNA. Proc Natl Acad Sci U S A 104:151–156

Guo X, Ernst NL, Carnes J, Stuart KD (2010) The zinc-fingers of KREPA3 are essential for the complete editing of mitochondrial mRNAs in *Trypanosoma brucei*. PLoS One 5:e8913

Hashimi H, Ziková A, Panigrahi AK, Stuart KD, Lukeš J (2008) TbRGG1, an essential protein involved in kinetoplastid RNA metabolism that is associated with a novel multiprotein complex. RNA 14:970–980

Hernandez A, Panigrahi A, Cifuentes-Rojas C, Sacharidou A, Stuart K, Cruz-Reyes J (2008) Determinants for association and guide RNA-directed endonuclease cleavage by purified RNA editing complexes from *Trypanosoma brucei*. J Mol Biol 381:35–48

Hernandez A, Madina BR, Ro K, Wohlschlegel JA, Willard B, Kinter MT, Cruz-Reyes J (2010) REH2 RNA helicase in kinetoplastid mitochondria: ribonucleoprotein complexes and essential motifs for unwinding and guide RNA (gRNA) binding. J Biol Chem 285:1220–1228

Jarmoskaite I, Russell R (2014) RNA helicase proteins as chaperones and remodelers. Annu Rev Biochem 83:697–725

Kafková L, Ammerman ML, Faktorová D, Fisk JC, Zimmer SL, Sobotka R, Read LK, Lukeš J, Hashimi H (2012) Functional characterization of two paralogs that are novel RNA binding proteins influencing mitochondrial transcripts of *Trypanosoma brucei*. RNA 18 (10):1846–1861. https://doi.org/10.1261/rna.033852.112

Kang X, Gao G, Rogers K, Falick AM, Zhou S, Simpson L (2006) Reconstitution of full-round uridine-deletion RNA editing with three recombinant proteins. Proc Natl Acad Sci U S A 103:13944–13949

Koslowsky D, Sun Y, Hindenach J, Theisen T, Lucas J (2013) The insect-phase gRNA transcriptome in *Trypanosoma brucei*. Nucleic Acids Res 42(3):1873–1886

Kumar V, Madina BR, Gulati S, Vashisht AA, Kanyumbu C, Pieters B, Shakir A, Wohlschlegel JA, Read LK, Mooers BH, Cruz-Reyes J (2016) REH2C helicase and GRBC subcomplexes may base pair through mRNA and small guide RNA in kinetoplastid editosomes. J Biol Chem 291:5753–5764

Law JA, O'Hearn SF, Sollner-Webb B (2008) *Trypanosoma brucei* RNA editing protein TbMP42 (band VI) is crucial for the endonucleolytic cleavages but not the subsequent steps of U-deletion and U-insertion. RNA 14:1187–1200

Madina BR, Kumar V, Metz R, Mooers BH, Bundschuh R, Cruz-Reyes J (2014) Native mitochondrial RNA-binding complexes in kinetoplastid RNA editing differ in guide RNA composition. RNA 20:1142–1152

Madina BR, Kumar V, Mooers BH, Cruz-Reyes J (2015) Native variants of the MRB1 complex exhibit specialized functions in kinetoplastid RNA editing. PLoS One 10:e0123441. https://doi.org/10.1371/journal.pone.0123441

McDermott SM, Stuart K (2017) The essential functions of KREPB4 are developmentally distinct and required for endonuclease association with editosomes. RNA 23:1672–1684

McDermott SM, Carnes J, Stuart K (2015a) Identification by random mutagenesis of functional domains in KREPB5 that differentially affect RNA editing between life cycle stages of *Trypanosoma brucei*. Mol Cell Biol 35:3945–3961

McDermott SM, Guo X, Carnes J, Stuart K (2015b) Differential editosome protein function between life cycle stages of *Trypanosoma brucei*. J Biol Chem 290:24914–24931

McDermott SM, Luo J, Carnes J, Ranish JA, Stuart K (2016) The architecture of *Trypanosoma brucei* editosomes. Proc Natl Acad Sci U S A 113:E6476–E6485

Panigrahi AK, Ziková A, Dalley RA, Acestor N, Ogata Y, Anupama A, Myler PJ, Stuart KD (2008) Mitochondrial complexes in *Trypanosoma brucei*: a novel complex and a unique oxidoreductase complex. Mol Cell Proteomics 7:534–545

Prabu JR, Muller M, Thomae AW, Schussler S, Bonneau F, Becker PB, Conti E (2015) Structure of the RNA helicase MLE reveals the molecular mechanisms for uridine specificity and RNA-ATP coupling. Mol Cell 60:487–499

Read LK, Lukeš J, Hashimi H (2015) Trypanosome RNA editing: the complexity of getting U in and taking U out. Wiley Interdiscip Rev RNA. https://doi.org/10.1002/wrna.1313

Reifur L, Yu LE, Cruz-Reyes J, Vanhartesvelt M, Koslowsky DJ (2010) The impact of mRNA structure on guide RNA targeting in kinetoplastid RNA editing. PLoS One 5:e12235

Robert-Paganin J, Réty S, Leulliot N (2015) Regulation of DEAH/RHA helicases by G-patch proteins. Biomed Res Int 2015:931857

Rusché LN, Cruz-Reyes J, Piller KJ, Sollner-Webb B (1997) Purification of a functional enzymatic editing complex from *Trypanosoma brucei* mitochondria. EMBO J 16:4069–4081

Sabatini RS, Adler BK, Madison-Antenucci S, McManus MT, Hajduk SL (1998) Biochemical methods for analysis of kinetoplastid RNA editing. Methods 15:15–26

Seiwert SD, Heidmann S, Stuart K (1996) Direct visualization of uridylate deletion in vitro suggests a mechanism for kinetoplastid RNA editing. Cell 84:831–841

Simpson RM, Bruno AE, Chen R, Lott K, Tylec BL, Bard JE, Sun Y, Buck MJ, Read LK (2017) Trypanosome RNA editing mediator complex proteins have distinct functions in gRNA utilization. Nucleic Acids Res 45:7965–7983

Stuart KD, Schnaufer A, Ernst NL, Panigrahi AK (2005) Complex management: RNA editing in trypanosomes. Trends Biochem Sci 30:97–105

Walbott H, Mouffok S, Capeyrou R, Lebaron S, Humbert O, van Tilbeurgh H, Henry Y, Leulliot N (2010) Prp43p contains a processive helicase structural architecture with a specific regulatory domain. EMBO J 29:2194–2204

Weng J, Aphasizheva I, Etheridge RD, Huang L, Wang X, Falick AM, Aphasizhev R (2008) Guide RNA-binding complex from mitochondria of trypanosomatids. Mol Cell 32:198–209

Ye P, Liu S, Zhu Y, Chen G, Gao G (2010) DEXH-box protein DHX30 is required for optimal function of the zinc-finger antiviral protein. Protein Cell 1:956–964

Chapter 6
Mitochondrial RNA Editing and Processing in Diplonemid Protists

Drahomíra Faktorová, Matus Valach, Binnypreet Kaur, Gertraud Burger, and Julius Lukeš

Abstract RNA editing and processing in the mitochondrion of *Diplonema papillatum* and other diplonemids are arguably the most complex processes of their kind described in any organelle so far. Prior to translation, each transcript has to be accurately trans-spliced from gene fragments encoded on different circular chromosomes. About half of the transcripts are massively edited by several types of substitution editing and addition of blocks of uridines. Comparative analysis of mitochondrial RNA processing among the three euglenozoan groups, diplonemids, kinetoplastids, and euglenids, highlights major differences between these lineages. Diplonemids remain poorly studied, yet they were recently shown to be extremely diverse and abundant in the ocean and hence are rapidly attracting increasing attention. It is therefore important to turn them into genetically tractable organisms, and we report here that they indeed have the potential to become such.

6.1 Introduction

6.1.1 General Overview

It is beyond reasonable doubt that the genome of all extant mitochondria is of bacterial origin and with high confidence derives from a single acquisition of an alpha-proteobacterium by an archaeal cell (Zimorski et al. 2014). The mitochondrial genome was then subject to progressive reduction by downsizing of the endosymbiont genome and via the transfer of genes into the nucleus and subsequent retargeting of their products into the organelle. This led to a stepwise conversion

D. Faktorová · B. Kaur · J. Lukeš (✉)
Institute of Parasitology, Biology Centre and Faculty of Sciences, University of South Bohemia, České Budějovice, Czech Republic
e-mail: jula@paru.cas.cz

M. Valach · G. Burger
Department of Biochemistry and Robert-Cedergren, Centre for Bioinformatics and Genomics, Université de Montréal, Montreal, Canada

J. Cruz-Reyes, M. Gray (eds.), *RNA Metabolism in Mitochondria*, Nucleic Acids and Molecular Biology 34, https://doi.org/10.1007/978-3-319-78190-7_6

of the endosymbiont into a mitochondrial organelle that is controlled largely from the nucleus (Lithgow and Schneider 2010; Gray 2012). Closest to the original proto-mitochondrial version seems to be the gene-rich mitochondrial genomes of jakobid flagellates, which belong to the supergroup Discoba (Burger et al. 2013). In several lineages, the gradual loss of genes resulted in a minimized genome containing just two protein-coding genes (Flegontov et al. 2015) or in a complete elimination of the mitochondrial genome (Maguire and Richards 2014). In other lineages that include both uni- and multicellular eukaryotes, organization of the mitochondrial genome acquired an almost limitless spectrum of forms and structures, which led some authors to postulate that "anything goes" in these organellar genomes (Burger et al. 2003). Recent research shows that this statement also applies to the expression of mitochondrial genes, as their transcripts are more often than not subject to diverse and complex forms of RNA editing, splicing, and processing.

Moreover, the structural and organizational diversity is not confined to the genome and transcriptome but also applies to the proteome of these organelles. Interestingly, only a minor fraction of proteins constituting the mitochondrion (= mitoproteome) is a remnant of the original alpha-proteobacterium, while most of them are of diverse prokaryotic (but other than alpha-proteobacterial) or eukaryotic origin (Szklarczyk and Huynen 2010). The evolution of the mitochondrial ribosome represents an illustrative example of numerous lineage-specific losses accompanied by gains of a substantial amount of novel proteins (Desmond et al. 2011). Since most of extant eukaryotic diversity is hidden in poorly studied protist lineages (Pawlowski et al. 2012), it is likely that their mitoproteomes will significantly differ from that of the prototypic ones in yeast and human. The mitoproteomes of these latter opisthokonts are by far the best studied and are at present the largest in terms of protein repertoire, as summarized in MitoCarta2.0 (Calvo et al. 2016). However, it seems that some protist mitoproteomes may be as complex as those of their multicellular relatives, as exemplified by the studies of the mitochondrion of *Acanthamoeba castellanii* (Gawryluk et al. 2014) and *Trypanosoma brucei* (Zíková et al. 2017).

T. brucei and related trypanosomatid flagellates contain a single canonical mitochondrion that generates ATP via oxidative phosphorylation, with oxygen being the terminal electron acceptor (Tielens and van Hellemond 2009; Škodová-Sveráková et al. 2015). It is likely that in terms of main metabolic setup, *Diplonema papillatum* (Fig. 6.1) and other diplonemids have a rather similar organelle (our unpublished data). This presumption and the relatedness with kinetoplastid flagellates indicate that the mitoproteome of diplonemids will be highly complex rather than reduced as is the case of disparate anaerobic or microaerophilic eukaryotes (Maguire and Richards 2014). The well-studied mitochondrion of *T. brucei* with over 1100 proteins (Dejung et al. 2016; Urbaniak et al. 2013; our unpublished data) is as complex as the mitochondrion of multicellular organisms. Moreover, its metabolism is highly adaptable to the drastically different environments of the insect vector and the bloodstream of the mammalian host (Verner et al. 2015). It is reasonable to assume that the mitochondrion of diplonemids (Fig. 6.1) will be more akin to the

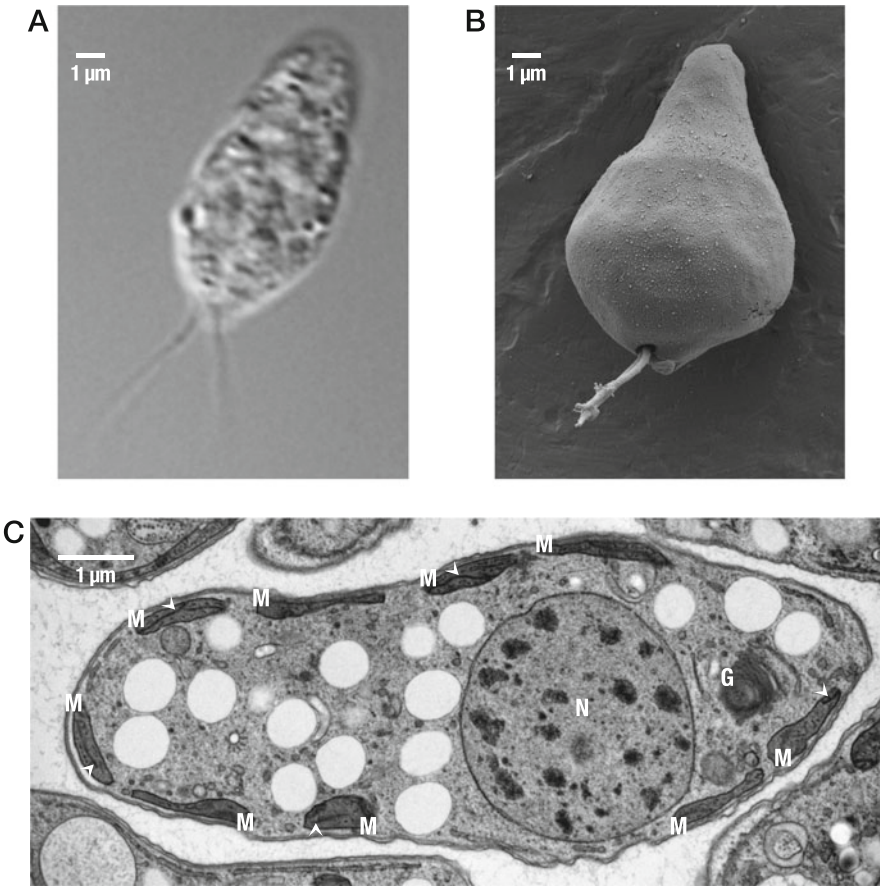

Fig. 6.1 Morphology of the model diplonemid, *Diplonema papillatum*. Light microscopy (**a**) and scanning electron microscopy (**b**) revealing the sac-like shape of the cell in culture and two heterodynamic flagella. (**c**) Transmission electron microscopy of a longitudinally sectioned cell with a prominent nucleus (N), single reticulated and peripherally located mitochondrion (M) with large discoidal cristae (arrowhead), and readily visible Golgi apparatus (G)

morphologically developed and metabolically highly active organelle of the insect-dwelling trypanosomes, especially since its nuclear genome is much larger (estimated at around 180 Mbp; our unpublished data) compared to that of the well-studied parasitic kinetoplastids (El-Sayed et al. 2005).

So far, diplonemids have been considered a marginal, rare, and rather insignificant group that received attention only thanks to its bizarre mitochondrial genome (see below). However, as they are emerging as major players in the world oceanic ecosystem, we predict that the era of diplonemids is just beginning.

6.1.2 Diplonemid Ecology, Taxonomy, and Phylogeny

With a single known exception (Triemer and Ott 1990), diplonemids seem to be confined to the marine environment including benthic waters. Yet in this largest planetary ecosystem, they are virtually omnipresent. In the frame of a global survey of marine microbial eukaryotes performed by the *Tara* Oceans expedition, based on the V9 region of the 18S ribosomal (r)RNA gene, over 85% of total eukaryotic plankton diversity is represented by unicellular eukaryotes (de Vargas et al. 2015). Diplonemids appeared among the most abundant groups, as they constitute the sixth most abundant (by reads of rRNA) and the third most diverse (by the number of operational taxonomic units, OTUs) eukaryotic group of the photic zone (de Vargas et al. 2015; Lukeš et al. 2015).

This came as a surprise since all the other prominently present eukaryotic groups were already well known, whereas diplonemids were until then considered rare and ecologically insignificant protists. In some stations of the *Tara* Oceans expedition, diplonemids reach up to 58% of all eukaryotes in the deeper mesopelagic zone (Flegontova et al. 2016) and were detected down to 6000 m in the poorly studied abyssopelagic zone (Eloe et al. 2011). Extensive sampling in the deeper pelagic layer, which is apparently the main habitat of diplonemids, further confirmed their prominent position among marine planktonic eukaryotes in terms of abundance and diversity (Flegontova et al. 2016).

The vast majority of marine diplonemids falls into a single clade dubbed the "deep-sea pelagic diplonemids" (DSPD) from deep oceanic environments (López-García et al. 2001, 2007; Lara et al. 2009) and was recently encountered at various depths ranging from surface to mesopelagic waters (Lukeš et al. 2015). The DSPD clade is also widespread in different geographical locations, ranging from tropical to polar regions, as well as from coastal to open ocean environments (Flegontova et al. 2016). Despite their diversity, ubiquity, and apparent abundance, we know close to nothing about the lifestyle, morphology, physiology, and biochemistry of the DSPD clade. Diplonemid species subjected to studies so far have been associated with parasitic or predatory lifestyles in plants, diatoms, and other marine protists (Schnepf 1994; Yabuki and Tame 2015). However, neither of the investigated species falls into the DSPD clade, which represents over 90% of diplonemid diversity.

The elusive DSPD diplonemids were, however, frequently encountered in a single-cell genomic survey of heterotrophic flagellates, conducted in the North Pacific Ocean (Gawryluk et al. 2016). Data generated from 10 individual cells, some of which belonged to OTUs most frequently represented in the *Tara* Oceans dataset, contain over 4000 protein-coding genes that fall into an ensemble of categories expected for heterotrophic protists. One striking feature is the high density of nonconventional introns that are absent from their kinetoplastid sister group (Gawryluk et al. 2016). Although we still have limited morphological and genetic information about the DSPD clade, it has now been formally described as a new class within Diplonemidea (Okamoto et al. 2018).

Moreover, significantly more information is available on the morphology, ultrastructure, and behavior of marine diplonemids not falling into the DSPD clade, but constituting several sister clades. These sac-like cells, highly variable in size and shape, have invariably two heterodynamic flagella inserted into a pronounced flagellar pocket and a DNA-rich mitochondrion with prominent lamellar cristae (Fig. 6.1). As is expected for a newly emerging speciose group of protists, the taxonomy and phylogeny of diplonemids is likely to evolve in the upcoming years.

6.1.3 Relationship of Diplonemids to Other Members of Euglenozoa

Diplonemids are part of the supergroup Euglenozoa, which includes two other morphologically and biochemically distinct main groups, kinetoplastids and euglenids (Adl et al. 2012; Cavalier-Smith 2016). This triumvirate was extended by the addition of anaerobic symbiontids (also called postgaardids) that were until recently placed among euglenids (Cavalier-Smith 2016). Symbiontids, which have a uniquely modified feeding apparatus and owe their name to their dependence on surface bacteria, are a poorly studied small group with only three genera described so far—*Postgaardia*, *Calkinsia*, and *Bihospites*. They were isolated from anoxic or low-oxygen environment, mainly from marine sediments (Yubuki et al. 2009, 2013; Breglia et al. 2010). Dependence on surface-bound episymbiotic bacteria along with hydrogenosome-like mitochondria with reduced cristae indicate a tight mutualistic relationship. Recently, symbiontids were shown to be present worldwide, similarly to the other euglenozoan groups, and they also seem to be more diverse than appreciated so far (Breglia et al. 2010; Edgcomb et al. 2011; Yubuki et al. 2013).

6.1.4 Mitochondrial Genome and Gene Structure

Despite the fact that all mitochondria are most likely derived from a single endosymbiotic event, mitochondrial genomes have evolved into myriad forms (Burger et al. 2003). The most diverse mitochondrial genomes are to be found among protists belonging to the supergroup Discoba (Smith and Keeling 2015). Jakobida harbor the most gene-rich mitochondrial genomes known (Burger et al. 2013), while anaerobic Metamonada exhibit mitochondrial reduction and even complete organelle loss (Karnkowska et al. 2016).

Arguably one of the most complex forms of mitochondrial DNA (mtDNA) evolved in diplonemid flagellates. *D. papillatum* carries in its organelle the largest amount of mtDNA known so far. The presence of an extraordinarily high amount of nucleic acids in its single mitochondrion was indicated by centrifugations of total

DNA in cesium chloride density gradients (Maslov et al. 1999). Later on, this observation was corroborated by staining mtDNA in situ, which revealed a strong continuous signal throughout the lumen of the reticulated organelle (Marande et al. 2005).

Flow cytometry experiments indicate that the *D. papillatum* nuclear genome has a size of about 180 Mbp (our unpublished data). In a more recent study, the cultured cells were stained simultaneously with an A + T-selective and nonselective dye, and the nuclear and mitochondrial signals were distinguished by color deconvolution, followed by quantification (Wheeler et al. 2012). This approach revealed massive inflation of the *D. papillatum* mtDNA, which with its estimated size of 270 Mbp not only exceeds that of the corresponding nuclear DNA but also represents the largest amount of DNA documented in any bacterium-derived organelle (Lukeš et al. unpublished). However, this enormous inflation does not reflect the gene content, which is rather ordinary, specifying subunits of respiratory complexes (six identified ORFs have unknown function) and the large and small subunit mitoribosomal rRNAs (Vlcek et al. 2011; Valach et al. 2014; Moreira et al. 2016).

Members of the genera *Diplonema* and *Rhynchopus*, as well as *Hemistasia phaeocysticola* have a multipartite mitochondrial genome (Vlcek et al. 2011; Yabuki et al. 2016). In *D. papillatum*, mtDNA is composed of thousands of non-interlocked circular chromosomes of at least 81 sequence classes that fall into two size categories—6 kb and 7 kb long, also labelled classes A and B, respectively (Marande et al. 2005) (Fig. 6.2). Within each class, chromosomes are essentially identical in sequence except for a short region called "cassette." Representing only about 5% of the chromosome, each cassette is composed of short unique 5′ and 3′ regions that flank a coding sequence, which is invariably a single gene fragment. With the sole exception of the small mitoribosomal rRNA, all genes are broken into up to 11 fragments, each of which resides on an individual chromosome (Valach et al. 2016); contiguous gene versions were not detected in mtDNA or nuclear DNA of *D. papillatum* (Figs. 6.2, 6.3 and 6.4). As a consequence of systematic fragmentation, not a single gene could be recognized at the outset of investigating the mitochondrial genome (Burger et al. 2016).

6.2 From Fragmented Genes to Contiguous Transcripts Via RNA Splicing

6.2.1 Splicing Types Found in Nature

As detailed above, genes in diplonemid mitochondria are systematically fragmented. However, mRNAs and rRNAs are, as usually, in one piece. Therefore, some kind of posttranscriptional mending must take place, which we have investigated mostly in *D. papillatum* and to some degree in *D. ambulator*, *Diplonema* sp. 2 [recently renamed to *Flectonema neradi* (Tashyreva et al. 2018)], and *Rhynchopus euleeides*.

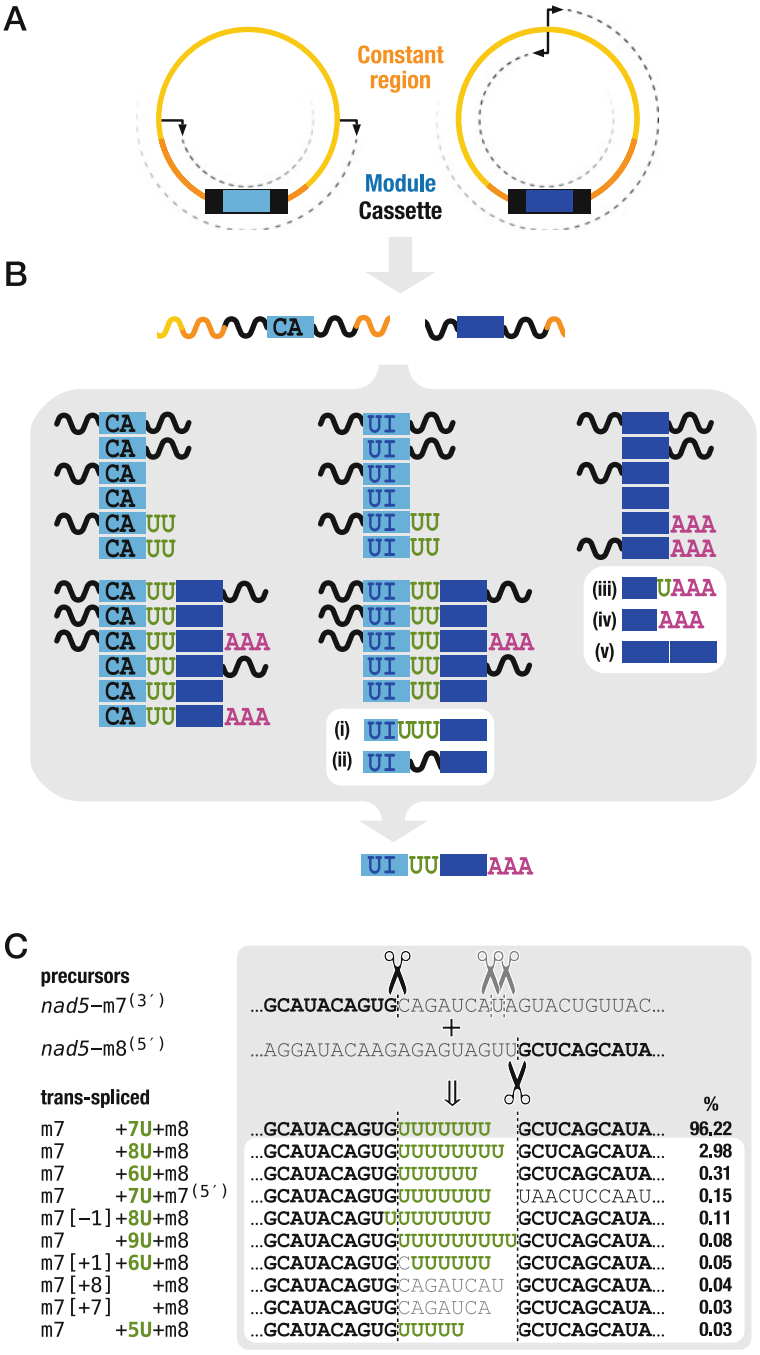

Fig. 6.2 Gene expression in diplonemid mitochondria. (**a**) Canonical circular mitochondrial chromosomes comprise a constant region of identical sequence across all members of a class

Our results reveal that the formation of contiguous mitochondrial mRNAs and rRNAs is diametrically different from conventional RNA splicing.

To summarize briefly, four major RNA splicing mechanisms exist across the various life forms and are classified according to the type of intervening sequence that is being eliminated: spliceosomal, tRNA (or archaeal), Group I, and Group II intron splicing (reviewed in Moreira et al. 2012). An additional less abundant type acting in fungi and vertebrates is IRE-mediated splicing that removes HAC1/XBP1 introns from pre-mRNA (Gonzalez et al. 1999). Each intron type is spliced by a distinct molecular machinery, be it a ribonucleoprotein complex (spliceosomal introns), catalytic RNA assisted by proteins (Groups I and II introns), or proteinaceous enzymes (tRNA and HAC1/XBP1 introns) (Hudson et al. 2015; Stahley and Strobel 2006; Zhao and Pyle 2017; Tanaka et al. 2011).

Initially, RNA splicing was viewed as an intramolecular (*cis*) reaction, removing an internal stretch of a pre-RNA and resealing adjacent exons. However, each of the abovementioned splicing types can also proceed in *trans*, i.e., the exons can reside on separate molecules, essentially representing halves of a pre-RNA broken apart within the intron.

6.2.2 RNA Processing Steps Prior to Trans-splicing in Diplonemid Mitochondria

Expression of fragmented genes in diplonemid mitochondria involves a unique mode of trans-splicing not seen before in any other system. The substrate for this particular trans-splicing is generated in a series of steps. First, gene pieces are

Fig. 6.2 (continued) [e.g., (**a**) and (**b**) in *D. papillatum*] and a unique cassette, which encloses a module (gene fragment). A cassette may be oriented in either sense relative to the constant region (illustrated at left and right). Long primary transcripts are initiated from the constant region by either two convergent promoters (left), or a bi-directional promoter (right), and extended into the other side of the constant region. (**b**) Separately transcribed single module precursors are processed in a highly parallelized process, which includes removal of 5′ and 3′ flanking noncoding regions, C-to-U, A-to-I, and U-appendage RNA editing of specific modules, 3′ polyadenylation of terminal modules, and trans-splicing of modules at processed ends (gray background). During the processing and trans-splicing, errors and their repair can take place: (i) exonucleolytic over-trimming of a module can be compensated for by a longer U-tract; (ii) 3′ flanking region of the upstream module can be retained instead of a U-tract; (iii) 3′ end over-trimming of can be compensated for by U-addition, even if the terminus is not normally a U-appendage site; (iv) polyadenylation of the terminal module may occur at over-trimmed sites; (v) two non-cognate modules can be joined together. Note that only the coding-strand transcripts are shown. (**c**) Examples of erroneous and error-compensating intermediates at the junction between the modules m7 and m8 of *nad5* detected in the total RNA from *Flectonema neradi* (*Diplonema* sp. 2). Coding and flanking noncoding regions are shown in black and gray, respectively. Note that the correctly processed, U-appendage-edited, and trans-spliced product represents the vast majority of detected RNAs

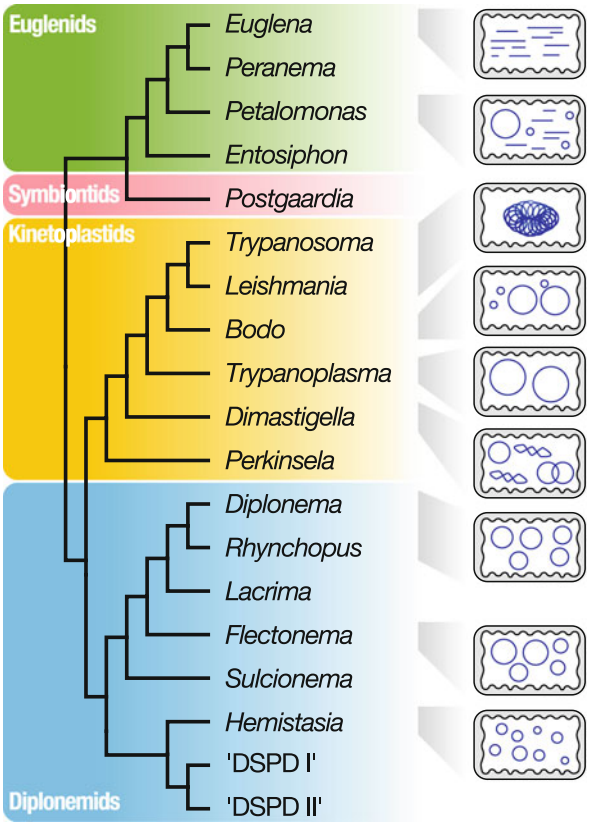

Fig. 6.3 Mitochondrial genome architectures and gene expression pathways in euglenozoans. Phylogenetic relationships among representative euglenozoan genera with their mitochondrial genome organization schematized. Euglenid mitochondria generally contain an assortment of linear molecules of variable length, though some species also harbor circular DNAs. Mitochondrial DNA of trypanosomatids, termed kinetoplast DNA (kDNA), is arranged into a single disc-shaped structure of catenated molecules. Bodonid species (*Bodo*, *Trypanoplasma*, and *Dimastigella*) contain non-catenated and relaxed or supercoiled circular molecules. Diplonemid mitochondrial circular chromosomes differ in size, with *Hemistasia* having particularly small chromosomes, as well as gene fragments

transcribed as long precursor molecules from a promoter located in the shared region of a mitochondrial chromosome. Although precise mapping of the transcription start site by in vitro capping experiments failed, the site was inferred to be located within the constant regions of chromosomes from precursor length determined by RNA circularization followed by RT-PCR across the ligation site (circRT-PCR) and amplicon sequencing (Kiethega et al. 2013) (Fig. 6.2a).

The promoter is most likely bi-directional (or two mirroring promoters exist in the constant region of circular chromosomes), since gene fragments are found encoded on either strand of the chromosome (plus and minus orientation of A-class and B-class

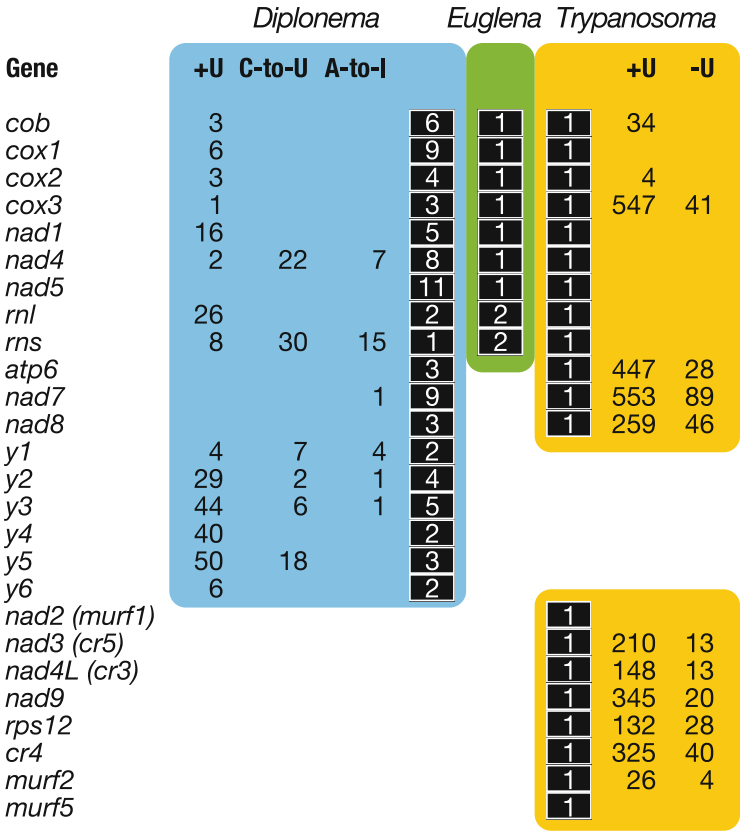

Fig. 6.4 Gene complement and editing site count across representative euglenozoans. Black rectangles indicate the presence of a gene (left column), with the number specifying the tally of precursor transcripts. Also shown is the total number of edits (+U, −U, C-to-U, A-to-I) in the corresponding mature transcript

chromosomes). In addition, antisense transcripts of individual gene fragments are detectable at low steady-state concentrations (Valach et al. 2014). Whether the amount of sense and antisense transcripts is regulated at the level of transcription initiation, transcription progressivity, or transcript degradation is currently not known.

The subsequent step in the expression of fragmented mitochondrial genes consists in end-processing of module transcripts. Processing intermediates, which are readily discernable by cDNA sequencing and circRT-PCR experiments, indicate that a combination of both endonucleolytic cuts and trimming are at work to generate transcripts that consist exclusively of coding regions (Fig. 6.2b). Only the 5′ ("first") module of protein-coding genes retains noncoding sequence, notably a 26- to 27-nt-long 5′ UTR (Kiethega et al. 2013).

Prior to trans-splicing, modules that will constitute the end of the mature transcripts undergo further maturation, notably addition of a homopolymer tail at the 3′

end. Transcripts of the "last" module from protein-coding genes are polyadenylated, forming the A-tail of mRNAs. Remarkably, A-tailed 3′-module transcripts belong to the most abundant precursors in total RNA, being present in certain cases (e.g., *cox1*) in a steady-state concentration comparable with that of the mature transcript (Marande and Burger 2007).

The last modules of both mito-rRNAs also receive a homopolymer tail. The large ribosomal subunit (mt-LSU) rRNA is polyadenylated. We reported previously that the transcript, once incorporated into the mitoribosome, has no A-tail (Valach et al. 2014). However, we realized recently that the result that led to this conclusion was due to an experimental artifact (see below "Limitations Encountered in Using the RNA-Seq Approach"). Reinvestigation of this issue by circRT-PCR demonstrates unambiguously that the A-tail length of mt-LSU rRNA (19–20 nt) remains unchanged after integration into the mitoribosome (Valach and Burger, unpublished data). The small ribosomal subunit (mt-SSU) rRNA from diplonemids studied so far is special in that its 3′ end carries a tail made from 8 Us. Curiously, in the kinetoplastid *Trypanosoma brucei*, both mt-rRNAs are modified by the addition of multiple terminal uridines (Adler et al. 1991).

Throughout eukaryotes, terminal adenylation or uridylation of rRNAs is generally a signal for degradation (Slomovic et al. 2010; Kuai et al. 2004). While exceptions to that rule have been reported for several taxa (Chaput et al. 2002; Mohanty and Kushner 2011), rigorous studies of either transcript stability or the state of rRNA actually incorporated into the ribosome are rare. Finally, prior to trans-splicing, certain modules will undergo RNA editing, which will be detailed in a later section.

6.2.3 *Succession of Posttranscriptional Processing Steps and Trans-splicing*

Contiguous mRNAs and mt-LSU rRNA of diplonemid mitochondria are formed through the joining of gene module transcripts that have been processed as described above. (Note that we use the term "module-transcript joining" synonymously with "trans-splicing") (Figs. 6.2, 6.3, 6.4 and 6.5). Intermediates of module-transcript end-processing, as well as trans-splicing, are readily detectable, not only in circRT-PCR experiments (Kiethega et al. 2013) and deep transcriptome sequencing (Moreira et al. 2016) but even in much less sensitive Northern hybridization (Marande and Burger 2007). This situation made *D. papillatum* an ideal system in which to investigate the temporal order of events.

Specifically, we observed a mixture of end-processing and trans-splicing intermediates, demonstrating that the succession of the individual posttranscriptional processing steps is not as strict as presented above (Fig. 6.2b). For example, module transcripts were detected that still carry adjacent, noncoding sequence at one terminus, while their other terminus is already trans-spliced to the neighbor module. This

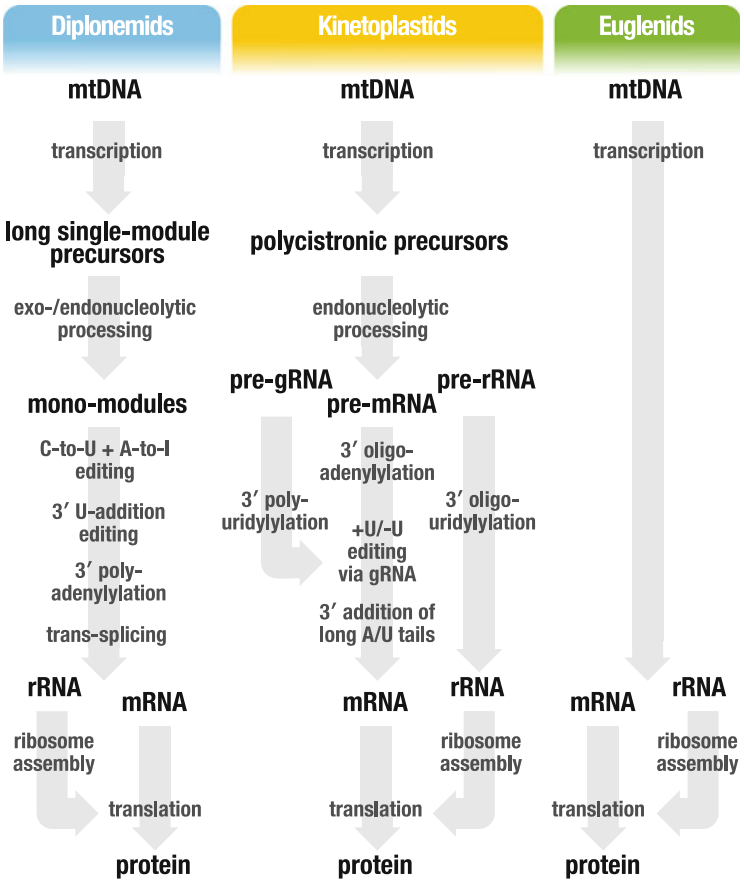

Fig. 6.5 Comparison of gene expression pathways among euglenozoans. Note that while the processes are sequential in kinetoplastids, diplonemids perform most processing and editing steps in parallel (see also Fig. 6.2)

shows that end-processing is not required to be completed for both module termini before trans-splicing can proceed. Further, polyadenylation of 3′-module transcripts is not a prerequisite for trans-splicing of their 5′ end to the upstream neighbor. Similarly, RNA editing of a module transcript via substitutions is not required to have taken place before trans-splicing. The only exception is U-appendage RNA editing. U-addition at module 3′ ends is completed before the corresponding terminus is joined to its downstream module or, in the case of terminal modules, before it is polyadenylated. Still, trans-splicing products with incompletely processed ends are the minority, as are those that are still pre-edited or not yet polyadenylated.

In summary, trans-splicing results in the correct sequential order of modules, yet proceeds without a particular directionality (e.g., 3′ to 5′). Thus transcript biogenesis in diplonemid mitochondria is a highly parallelized process (Kiethega et al. 2013).

6.2.4 Partner Selection in Trans-splicing

In diplonemids of the D/R clade, about 80 mitochondrial module transcripts have to be trans-spliced to their correct partner, raising the question how cognate module recognition is achieved. Cis sequence elements such as those adjacent to trans-splicing sites of conventional introns are not discernable, nor are conserved primary or secondary structure elements that are shared by all splice sites (Kiethega et al. 2011). Therefore, we posit trans-acting factors that recognize module transcripts to be joined and align them tail to head for trans-splicing.

Kinetoplastids possess guide RNAs, an abundant species of ~50-nt-long transcripts with a 5′-triphosphate and a U-tail, which are involved in mitochondrial uridine insertion and deletion RNA editing (Aphasizhev and Aphasizheva 2011; Read et al. 2016). We speculated initially that such molecules might guide trans-splicing in diplonemids, yet *Diplonema* mitochondria do not contain such an RNA species (Kiethega et al. 2013). Another conceivable kind of trans-acting splice guides would be full-length antisense mRNAs and antisense rRNA, serving as a single template for all splice junctions of a given gene. Since full-length genes are not present in *Diplonema* nuclear or mitochondrial DNA, these antisense transcripts would have to be produced by an RNA-dependent RNA polymerase (Valach et al. 2014), using sense transcripts as a template. Alternatively, there might be multiple (i.e., a total of 69) short antisense RNAs, each complementary to a single module junction.

We tested the "antisense RNA hypothesis" for *cox1* and *rnl*. For *cox1*, we performed exhaustive in silico analyses in an attempt to detect potential splice guides. Indeed, for each of the junctions, sequences were identified in the mitochondrial and nuclear genome that have the potential to be transcribed into splice guides (Kiethega et al. 2011); in turn, RT-PCR experiments indicated the existence of splice guides for five of the eight junctions (Kiethega et al. 2013). For *rnl*, which is ~100× more highly expressed, RT-PCR returned a readily discernable antisense product, Northern experiments showed a weak and smeary signal, and deep sequencing of a stranded cDNA library made from total RNA using an approach, which produces di-tagged first-strand cDNAs (ScriptSeq kit), yielded ~2.5% read coverage of the complementary strand bridging the *rnl*-m1/*rnl*-m2 junction. This rate is more than two times above the 1% of spurious antisense reads considered typical for the methodology (Valach et al. 2014).

Yet, these results must be considered with caution. The RT-PCR technique may produce artifactual antisense products, e.g., by polymerase template switching. Moreover, in more recent RNA-Seq experiments (Valach et al., unpublished data), we noted a considerable variation in the depth of junction-crossing antisense reads between libraries made from different RNA preparations (1–6%). Furthermore, total RNA-Seq libraries, made using the first-strand dUTP-cDNA approach (as implemented in the TruSeq kit) (Parkhomchuk et al. 2009) of *D. ambulator*, *F. neradi* (*Diplonema* sp. 2), and *R. euleeides*, showed only a coverage of 1% (i.e., background level). In sum, at the current time, it is uncertain if *Diplonema* cells

indeed produce a significant steady-state level of genuine junction-crossing antisense RNAs for *cox1* or mt-LSU rRNA.

Instead of RNA guides, trans-splicing could also be directed by guide proteins. Such proteins must be capable of binding selectively to specific RNA sequence motifs. Sequence-specific RNA-binding proteins are generally composed of several conserved RNA-binding domains that engage in base-dependent interactions with RNA and form a three-dimensional shape that is complementary to that of the recognized RNA motif (Ban et al. 2015). The most common and best-studied RNA-binding proteins are characterized by either tristetraprolin (TTP)-type tandem zinc finger domains, pentatricopeptide repeat protein (PPR) domains, Pumilio-FBF (Puf) domains, or RNA recognition motif (RRM) domains. We detected genes from the three latter families in the preliminary version of the *D. papillatum* nuclear genome sequence (our unpublished data). It remains to be confirmed, in silico and experimentally, which of these predicted proteins are located in the mitochondrion.

6.2.5 *Accuracy of Module Trans-splicing*

With several dozen distinct gene module transcripts in the diplonemid mitochondrion, what is the trans-splicing accuracy? Deep transcriptome sequencing of total *D. papillatum* RNA shows on average ~0.1% mis-spliced transcripts, with certain modules being considerably more "promiscuous" than others (Fig. 6.2c). For example, in a total RNA library, as much as ~16% of trans-spliced *cox1*-m7 3′ termini have been joined incorrectly, i.e., predominantly to *cox1*-m6 instead of *cox1*-m8. In contrast, poly-A libraries contain only about 0.4% of mis-joined *cox1*-m7 products. Thus, incorrectly joined modules appear to be eliminated by some quality control mechanism in mature polyadenylated mRNAs. Mis-splicing might be caused by short identical sequence motifs. A preliminary search (≥6-nt-long motifs within 20-nt from the junction) did not reveal recurrent patterns. The analysis has to be extended to more distant regions and also consider secondary structure motifs.

Interestingly, a recent investigation of the distantly related diplonemid *H. phaeocysticola* (Yabuki et al. 2016) recovered rare cases of mitochondrial transcripts in which the first *cox1* module was joined to downstream modules other than the expected module 2. These findings were interpreted as indicative of an mRNA assembly pathway containing a step of module-transcript insertion in contrast to a "concatenation" model described for *D. papillatum*. Although this suggestion is an intriguing possibility that merits further study, in the light of the existence of module mis-joining in all D/R diplonemids studied thus far (Valach et al. 2016), it seems more plausible that the rare transcripts with unexpected module order in *Hemistasia* also represent dead-end intermediates.

6.2.6 Speculations on the Trans-splicing Reaction and Machinery

While the process of trans-splicing in diplonemid mitochondria is quite well characterized, open questions remain about the reaction itself. Given the absence of conserved nucleotides at the splice junctions, a ribozyme reaction mechanism is unlikely, thus favoring the hypothesis of an enzyme-based ligation of module transcripts. For example, splicing of conventional tRNA introns and of HAC/XBP1 involves an end-joining reaction catalyzed by RNA ligases of the T4 Rnl or the RtcB family (Popow et al. 2012). Preliminary analyses of the nuclear genome draft from *D. papillatum* show that it encodes proteins of the RtcB family. Some family members will be involved in the splicing of nuclear tRNA introns, while others might join mitochondrial module transcripts. We postulate that module ligation and matchmaking are performed by an integrated molecular machinery—the hypothetical joinosome (Valach et al. 2016)—whose identification is our priority.

Interestingly, a second case of unorthodox trans-splicing has been reported in mitochondria of certain dinoflagellates. One of the mitochondrion-encoded gene, *cox3*, is broken up into two separate pieces, while its transcript is contiguous (Jackson and Waller 2013). Whether the machinery involved shares communalities with the system in diplonemids remains to be investigated.

6.2.7 Limitations Encountered in Using the RNA-Seq Approach

By investigating the mitochondrial transcriptome of diplonemids, we became aware of several limitations of the RNA-Seq approach (see also Ozsolak and Milos 2011; Levin et al. 2010). One problem is that read coverage only partially represents the actual steady-state level of a transcript, especially when the library construction protocol, as in our case, uses hexamer primers for initiating first-strand cDNA synthesis. Not only does coverage drop strongly toward the template's extremities, but also internally drastic fluctuations occur, probably due to differences in efficiency of primer annealing to particular sequence contexts. An important challenge in our analyses was the low read coverage in homopolymer tracts, likely caused by inefficient progressivity at the stage of reverse transcription and sequencing.

Strand specificity is another issue, especially when analyzing the level of genuine antisense transcripts. We noted that the degree of spurious antisense reads depends on the sequence of the gene in question and may be above or below the overall vendor-stated rate of a given library construction protocol. To determine exactly the level of spurious antisense products, controls with an in vitro-synthesized RNA should be performed so that the portion of genuine antisense transcripts in the sample can be reliably assessed. Our approach was to synthesize in vitro an ~200-nt-long RNA that covers the *rnl*-m1/*rnl*-m2 junction and—for cost reasons—mix it into an

RNA preparation of another organism to be sequenced, construct a stranded RNA-Seq library of this mix, and sequence it.

Further, we encountered the problem that capture probes used for eliminating over-abundant transcripts, such as rRNAs, are not always removed completely from the sample prior to library construction. Remnants of the capture probe prime reverse transcriptase during first-strand synthesis, generating artificially profuse amounts of reads all starting at the same position. Capture probes are biotinylated for easy removal with streptavidin-coated magnetic beads after annealing with their target rRNA. We assume that sample contamination occurred because of a too low ratio of beads to capture probe and/or because of incomplete biotinylation of the oligonucleotide.

A final unexpected issue was that different RNA-Seq library construction kits are not equally effective in reproducing A-tails. Control experiments with circRT-PCR confirmed that mt-LSU rRNA has indeed 19–20 As at its 3′ end, just as determined via the reads from the ScriptSeq library (Valach et al. 2014), while the TruSeq library returned only 0–2-nt-long A-tails. We assume that the particular mix of random primers used by the TruSeq protocol is biased against annealing with A-tracts.

6.3 From Defective to Functional Products Via RNA Editing

As alluded to in the previous section, the convoluted mitochondrial gene expression in diplonemids does not stop at ribonucleolytic processing and covalent joining to form the functional mRNA or rRNA. Certain module transcripts undergo additional maturation steps, which result in nucleotide-level changes of the transcript sequence.

6.3.1 Types of RNA Editing Systems in Mitochondria

In general, RNA modifications corresponding to nucleotide insertions, deletions, or substitutions are referred to as RNA editing (reviewed in Knoop 2010). They may take place directly during transcription or at later maturation stages, may involve a variety of enzymatic activities (e.g., base deaminases, nucleases, ligases, 3′ or 5′ polymerases), and may affect mRNAs, rRNAs, or tRNAs, as well as other types of transcripts like miRNAs, ncRNAs, or retrotransposons (reviewed in Knoop 2010; Nishikura 2016). We first briefly overview the diversity of RNA editing mechanisms in mitochondria (Table 6.1), with emphasis on five instances where the enzymatic players have been characterized, before addressing the peculiarities of the diplonemid RNA editing.

C-to-U substitution is commonly encountered in land plant organelles, with hundreds to thousands of editing events per genome (reviewed in Takenaka et al.

Table 6.1 Diversity and distribution of RNA editing types in mitochondria

Type of change		Distribution	Transcript category	Selected references
Substitution	C-to-U	Land plants	mRNA	Reviewed in Takenaka et al. (2013)
		Slime molds	mRNA	Bundschuh et al. (2011)
		Heteroloboseans	mRNA	Rüdinger et al. (2011) and Fu et al. (2014)
		Diplonemids	mRNA, rRNA	Moreira et al. (2016)
		Malawimonads	mRNA	Authors' unpublished data
	U-to-C	Land plants	mRNA	Reviewed in Takenaka et al. (2013)
	A-to-I	Diplonemids	mRNA, rRNA	Moreira et al. (2016)
	Various	Dinoflagellates	mRNA, rRNA	Lin et al. (2002) and Jackson et al. (2007)
Insertion	Predominantly C (also U, A, G)	Slime molds	mRNA, rRNA, tRNA	Bundschuh et al. (2011), Mahendran et al. (1991) and Chen et al. (2012)
	Predominantly G (also A, C, U)	Heteroloboseans	mRNA, rRNA, tRNA	Yang et al. (2017)
	A	Dinoflagellates	mRNA, rRNA	Jackson et al. (2007)and Jackson and Waller (2013)
		Metazoans	tRNA	Yokobori and Pääbo (1995)
	U	Metazoans	mRNA	Vanfleteren and Vierstraete (1999) and Lavrov et al. (2016)
		Kinetoplastids	mRNA	Reviewed in Read et al. (2016)
		Diplonemids	mRNA, rRNA	Moreira et al. (2016)
	Various (5′ end)	Amoebozoans	tRNA	Jackman et al. (2012)
		Fungi	tRNA	Laforest et al. (1997)
		Heteroloboseans	tRNA	Authors' unpublished data
	Various (3′ end)	Jakobids	tRNA	Leigh and Lang (2004)
		Metazoans	tRNA	Segovia et al. (2011)
Deletion	A	Slime molds	mRNA	Gott et al. (2005)
	U	Kinetoplastids	mRNA	Reviewed in Read et al. (2016)

2013; see also Chap. 9). Although the enzyme responsible for the deamination reaction has not yet been unambiguously identified, a plethora of ancillary cofactors has been catalogued (reviewed in Sun et al. 2016) and the indirect experimental evidence has been converging on the DYW family of PPR proteins as the catalytic component (Salone et al. 2007; Shikanai 2015).

A different process takes place in the mitochondria of slime molds such as *Physarum polycephalum* (Bundschuh et al. 2011; Mahendran et al. 1991), where the plentiful mono- and dinucleotide insertions at internal sites in mitochondrial transcripts occur co-transcriptionally, probably relying on the interplay between the RNA polymerase complex and its substrate DNA (Visomirski-Robic and Gott 1997; see also Chap. 8). However, the exact mechanism of this system remains to be elucidated. Much better understood is the editing of mt tRNA at their 5′ and 3′ ends, a posttranscriptional nucleotide insertion process observed in various eukaryotic clades (Table 6.1). Several amoebozoans replace 5′ terminal nucleotides of their mt tRNA employing an unconventional 3′ to 5′ polymerase of the Thg1 family (Abad et al. 2011; see also Chap. 7). Editing of tRNA at 3′ end can proceed via polyadenylation by a 3′ terminal adenylyltransferase (poly-A polymerase), generating a missing secondary structure element (reviewed in Rammelt and Rossmanith 2016).

Finally, one of the best understood RNA editing processes takes place in the mitochondrion of kinetoplastids (Benne et al. 1986), where most mRNAs undergo extensive insertion and/or deletion of U residues by a complex ribonucleoprotein machinery (reviewed in Read et al. 2016; see also Chap. 5). The multicomponent editosome includes endonuclease, U-specific exoribonuclease, terminal uridylyltransferase (TUTase), and ligase activities, which for each edited site complete a cycle consisting of cleaving the transcript, inserting/deleting a number of Us specified by a partially complementary guide RNA, and religating the broken strand.

6.3.2 Idiosyncratic RNA Editing in Diplonemid Mitochondria

6.3.2.1 Appendage of Uridines

RNA editing in *D. papillatum* mitochondria was noted early on in the *cox1* cDNA, which contained six nonencoded Ts inserted between its modules 4 and 5 (Marande and Burger 2007). Once high-throughput cDNA sequencing technologies made possible a comprehensive investigation of RNA editing sites, a more complex picture emerged: in this diplonemid, 240 Us are inserted at 18 sites distributed across 14 out of its 18 genes (Moreira et al. 2016) (Figs. 6.2b, c and 6.4). Insertions of no other nucleotide besides U, nor nucleotide deletions, have been detected in any diplonemid analyzed to date.

While many U-tracts are shorter than the one in *cox1*, a stretch of as many as 26 Us is added in the middle of the mt-LSU rRNA (*rnl*) (Valach et al. 2014). Recently, we have confirmed the presence of even more impressive 50 Us in a row in the mature transcript of the (unassigned) gene *y5* (Valach et al. 2017). Such long U-tracts blur the line between the conventional definition of RNA editing (a single or a couple of affected nucleotides at a single site) and posttranscriptional modifications traditionally not considered to represent RNA editing, such as terminal polyuridylylation.

In *Diplonema*, the U residues are not inserted in a cut-add-reseal strategy as in kinetoplastid RNA editing. Instead, they are appended to 3′ termini of processed

modules prior to trans-splicing. First, all identified U insertions are confined to module junctions or to 3′ ends of last modules, just upstream of poly-A tails in case of mRNAs. Second, circular RT-PCR, 5′ and 3′ RACE, and primer extension assays showed that in *D. papillatum*, no *cox1* mRNA trans-splicing intermediate lacks the six Us after the modules 4 and 5 have been joined together, nor does it contain the six Us attached to the downstream module 5. It is exclusively the 3′ end of the upstream module 4 to which the U-tract is appended (Kiethega et al. 2013). Comprehensive investigation of the entire transcriptome further confirmed that only the 3′ end-processed module transcripts are uridylylated, irrespective of the maturation state of the 5′ end of that same module (Moreira et al. 2016). In this respect, the U-appendage pathway is similar to trans-splicing, which can also proceed even if the opposite end of a module that does not participate in module joining is incompletely processed (see the previous section; Fig. 6.2b).

The close relationship between module transcript joining and uridylylation has been further corroborated by our deep-coverage transcriptome data from *D. papillatum* and three additional diplonemids, revealing transient errors or "background noise." At a frequency around 0.1%, Us (mostly 1 to 3) are being added even at module 3′ ends that normally do not undergo U-appendage RNA editing. Interestingly, the vast majority of these abnormal U-addition events occur in trans-spliced transcripts whose upstream partner's 3′ end is several nucleotides shorter, with the Us compensating for the missing sequence (Valach et al. 2017). We thus hypothesize that the same process ensuring the usual U-appendage RNA editing can also repair a deletion at a module junction, which could have arisen from erroneous over-trimming during module transcript end-processing (Fig. 6.2b, c). Curiously, at certain, but not all, junctions usually separated by a U-tract (e.g., *nad5*-m7/m8 in *F. neradi* [= *Diplonema* sp. 2]), we also observe rare (<1%) occurrences of two cognate modules being joined together without the U-tract (Valach et al. 2017). However, in these cases, the missing sequence is compensated by a sequence stretch originating from the upstream module's 3′ flanking region, which is present instead of the expected U-tract (Fig. 6.2b, c). It remains to be seen whether these defective trans-spliced products are translated or rather are discarded by some downstream control mechanism, as is apparently the case for mis-joined, non-cognate modules (see Sect. 2). In any case, these two observations—gap filling by U-addition or partial retention of a 3′ flanking region—imply that some kind of a molecular ruler measures the length of the module transcripts or the distance between the two RNA ends to be joined. Likely candidates are the factors involved in junction recognition (see below).

6.3.2.2 Clustered Substitutions of Adenosines and Cytidines

The screening for cDNA vs. genome differences further unveiled 85 cytidine-to-uridine (C-to-U) substitutions, well known from organelles of many species (Table 6.1). In addition, we discovered 29 adenosine-to-guanosine (A-to-G) substitutions in half of the *D. papillatum* genes (Moreira et al. 2016) (Fig. 6.4). These substitutions indicate C-to-U and A-to-I base deamination (inosine is read as

guanosine during reverse transcription). Indeed, A-to-I deamination could readily be demonstrated experimentally (Moreira et al. 2016). While this type of deamination is common for tRNAs, ours was the first report of its kind for mitochondrial mRNAs and rRNAs. Diplonemid mitochondria also show an exceptionally high rate (>95%) of RNA editing at a given site, and further, in most instances, diplonemid editing sites congregated in clusters denser even than those of the so-called hyper-edited segments in metazoan nuclear transcripts (Wahlstedt and Ohman 2011) (Fig. 6.4).

The latter two features are particularly intriguing. As a general rule, we considered as a cluster a group of adjacent sites where more than half of the potentially editable residues (As + Cs) in a row were indeed edited. In *D. papillatum*, in all but one cluster (*y5*-m1), every single C in a cluster is edited, as are most As (Moreira et al. 2016). For example, in an 85 nt-long region of mt-SSU rRNA, all 15 As and all 30 Cs are substituted. Although most sites are edited to high levels, there are few partially edited (5–40% rate) sites, with editing rates generally slightly higher for C-to-U than for A-to-I substitutions. Still, all of these occur within a cluster or at its boundaries and thus may indicate "misfiring" of the editing enzyme(s).

Our comprehensive analyses of trans-splicing and editing intermediates in *D. papillatum* also revealed that substitution RNA editing in a cluster progresses stochastically and not directionally. As mentioned in the previous section on trans-splicing, substitution editing is essentially completed before trans-splicing begins; pre-edited or partially edited module transcripts that are already trans-spliced are found only at below 5% (Moreira et al. 2016) (Fig. 6.2).

6.3.3 *Functional Consequences of RNA Editing*

Both types of RNA editing in diplonemid mitochondria appear to be critical for the function of the affected transcripts. For example, in the case of *rnl*, the long U-tract is predicted to form segments of two helices of the mt-LSU's central domain 0 (Valach et al. 2014). The six Us of the *cox1* mRNA add codons for amino acids that restore the three-dimensional structure of the protein (Kiethega et al. 2011), whereas the two Us of the *nad4* transcript rectify the reading frame of the coding sequence (Moreira et al. 2016). In several mature transcripts (e.g., *cox3*, *y3*), U-appendage together with polyadenylation creates the termination codon, and in *nad1* mRNA, the 16 nt-long U- tract at its 3′ end adds codons for five additional phenylalanyl residues to the polypeptide, thus completing the C-terminal membrane-spanning helix. Similarly, substitution RNA editing of *nad4* mRNA leads to a protein that contains all its hydrophobic transmembrane helices instead of lacking the second helix (Moreira et al. 2016). Comparative analysis of the gene across four diplonemid species demonstrated that the proteins encoded by edited mRNAs became more similar to one another, as well as to homologs from other organisms.

Interestingly, dense C-to-U and A-to-I RNA editing results in codons rich in U and G (I) residues, which mostly specify apolar amino acids. In addition,

uridylylation creates UUU codons, which code for the hydrophobic phenylalanine residue. Apolar and hydrophobic amino acids being favored in membrane-embedded or membrane-anchored proteins, one can easily imagine how these two types of RNA editing in particular could become evolutionarily fixed for mending the deterioration of diplonemid genes, which all encode proteins of this class.

6.3.4 Predicted Components of the Editing Machineries

Based on our insights into diplonemid mitochondrial RNA editing described above, we have attractive working hypotheses about the nature of the enzymes involved in the two types of RNA editing. Akin to kinetoplastids, diplonemids add Us at the 3′ end of mitochondrial transcripts, suggesting that U-appendage RNA editing is performed by a TUTase enzyme similar to RET2 of the trypanosome editosome. For substitution RNA editing, a nucleotide/base excision-replacement system is conceivable, but our current data rather indicate that the mechanism relies on base deamination. Since the C-to-U and A-to-I edits are closely spaced and display no ordering of pre-edited and edited positions in transcript intermediates, we speculate that an enzyme able to deaminate both Cs and As is involved. Interestingly, a precedent for such an enzyme was discovered in the kinetoplastid *T. brucei* (Rubio et al. 2007). According to our preliminary analyses, several genes potentially encoding proteins with a nucleotidyltransferase or deaminase domain are present in the draft nuclear genome of *D. papillatum* (our unpublished data).

As in the case of trans-splicing (Kiethega et al. 2011), no cis-elements have been identified in the genome sequence that have the potential to direct the enzymatic machinery to the RNA editing sites (Moreira et al. 2016). This led us to postulate that all three processes—the trans-splicing, U-appendage, and substitution RNA editing—are guided by trans-acting factors (Valach et al. 2016). Among the numerous RNA-binding protein (RBP) families that were mentioned in the previous section and that could be implicated in mitochondrial RNA processing in *D. papillatum*, PPR proteins have emerged as primary candidates. They are not only the most notable cofactors of C-to-U editing in land plant organelles (Sun et al. 2016) but also serve as cofactors of numerous other organellar RNA transactions in a wide variety of organisms (Manna 2015).

6.4 Comparison of Mitochondrial Gene Expression Across Euglenozoa

In molecular biology textbooks, expression of genetic information is simple and straightforward. However, in some organisms it is surprisingly derived, incomprehensible, and gratuitously inefficient. This applies not only to diplonemids but also

to those protists from other euglenozoan groups. Since no molecular data are currently available about symbiontids, we will compare the expression of mitochondrial genes among the three other euglenozoan groups—diplonemids (besides *Diplonema papillatum* also represented by *Diplonema ambulator, Flectonema neradi* (*D.* sp.2), and *Rhynchopus euleides*), trypanosomatids (represented by *Trypanosoma brucei*), and euglenids (represented *by Euglena gracilis*, *Peranema trichophorum*, and *Petalomonas cantuscygni*). What is currently known about the organization of their mtDNA and about the mitochondrial gene expression of *D. papillatum*, *T. brucei*, and *E. gracilis* is summarized in Figs. 6.3, 6.4 and 6.5.

6.4.1 *Mitochondrial A + T Content and Gene Complement Throughout Euglenozoa*

All euglenozoans carry a single mitochondrion with discoidal cristae, with possibly the only exception being the euglenid *P. trichophorum*, which possesses several small elongated mitochondria (Roy et al. 2007). While packaging of mtDNA into a dense single kinetoplast remains a character exclusive to kinetoplastids, mtDNA in diplonemids and euglenids is homogenously distributed throughout the organellar lumen and is only exceptionally organized into tiny bodies or foci. The A + T content of mtDNA varies across euglenozoans—it has a typically higher A + T content in *T. brucei*, *R. euleides, E. gracilis*, and *P. cantuscygni*, but in *P. trichophorum* and *D. papillatum*, the A + T content is unusually low (Roy et al. 2007; Dobáková et al. 2015).

Regardless of its structure (Fig. 6.3), the mitochondrial genome of euglenozoans has a very similar gene composition. It is typically composed of subunits of four respiratory complexes, complex I (NADH dehydrogenase; *nad* genes), complex III (ubiquinone-cytochrome *c* oxidoreductase; gene *cob*), complex IV (cytochrome *c* oxidase; *cox* genes), and complex V (ATP synthase; gene *atp6*), and two mitoribosomal RNAs (*rnl* and *rns*) (Faktorová et al. 2016) (Fig. 6.4). Moreover, mtDNA in *T. brucei* also encodes ribosomal protein Rps12 (Alfonzo et al. 1997). No tRNA genes have been identified in any euglenozoan mitochondrial genome and therefore have to be imported from the cytoplasm (Alfonzo and Söll 2009).

6.4.2 *Comparison of* D. papillatum *Genome Structure with Other Diplonemids*

Diplonemid species studied to date possess several classes of circular chromosomes. Compared to the 6.0 kb and 7.0 kb classes in the case of *D. papillatum*, the sizes in the other studied species vary from 4.5 kbp to >6.7 kbp (with a majority at ~5 kbp) in *D. ambulator*, from ~5 kbp to ~10 kbp in *Diplonema* sp. 2 (= *F. neradi*;

Tashyreva et al. 2018), and from 5 kbp to 12 kbp (with a majority at ~7 kb and ~8 kb) in *R. euleeides* (Kiethega et al. 2011; Valach et al. 2017). In *D. papillatum* almost every gene split into fragments (up to 11) and each piece is encoded on a separate chromosome (Moreira et al. 2016) (Fig. 6.4). While the fragmentation pattern is essentially identical in the other investigated D/R clade species, up to eight gene pieces were found to be encoded on the same chromosome (Valach et al. 2017).

In *H. phaeocysticola*, the size of mitochondrial chromosomes sequenced so far is significantly smaller (2.7–3.2 kb), with twice as many half-sized gene fragments (Yabuki et al. 2016), and a similar situation seems to be the case in the newly isolated species belonging to the same clade (our unpublished data). Moreover, currently we are trying to shed more light on this group by studying the mitochondrial genome structure of newly described diplonemid species that belong to the genus *Rhynchopus* (*Rhynchopus humris* and *Rhynchopus serpens*) or to the newly described environmental clade (*Lacrimia lanifica*), and even a novel early-branching clade, represented by *Sulcionema specki* (Tashyreva et al. 2018).

6.4.3 Kinetoplastids: Uridines In and Out

Kinetoplastids are either free-living (e.g., *Bodo saltans*) or parasitic protists, which include human parasites of major medical importance, such as members of the genera *Trypanosoma* and *Leishmania*. They are characterized by a kinetoplast, a compact mass of mtDNA composed of dozens of maxicircles and thousands of minicircles (Shapiro and Englund 1995; Stuart and Feagin 1992). Maxicircles (~20 kbp) represent functional equivalents of mtDNA in other organisms. Most of the mitochondrial genes (12 out of 18) are literally encrypted (Fig. 6.4). This means that their transcripts have to undergo a process of RNA editing, which restores meaningful open reading frames that are translatable (Fig. 6.5). Since its first description in *T. brucei* (Benne et al. 1986), many distinct and unrelated types of RNA editing have been described in organisms across the entire tree of life (Read et al. 2016). In *T. brucei* and other kinetoplastid protists, RNA editing is guided by small minicircle-encoded molecules called guide RNAs (gRNAs) that serve as template for the insertions and/or deletions of uridines into the pre-edited sequence at specific positions (Aphasizhev and Aphasizheva 2011).

Interestingly, about a thousand distinct minicircle-encoded gRNAs, together with more than 70 different nucleus-encoded proteins, are necessary for proper expression of the small complement of 18 mitochondrion-encoded genes (Alfonzo et al. 1997; Verner et al. 2015; Read et al. 2016). More specifically, in addition to well-described RNA editing core complex (RECC) or the 20S editosome (Göringer 2012), several other ribonucleoprotein complexes, e.g., the MRB1 complex, were recently shown to be involved in the RNA editing and processing machinery (Ammerman et al. 2012; Read et al. 2016; Dixit et al. 2017).

The uridine insertion/deletion type of RNA editing in kinetoplastids, and even more the obscure and still unrecognized machinery for trans-splicing associated with uridine insertions and cytidine-to-uridine and adenine-to-inosine substitution RNA editing in diplonemids, appears extremely costly in comparison to their benefits. So far, no advantages of these strategies have been proposed, leading to the speculation that they most likely originated as a result of constructive neutral evolution (Flegontov et al. 2011; Lukeš et al. 2011).

6.4.4 Euglenids: Surprises in Their Own Right

It was hoped that elucidation of the structure and expression of mitochondrial genome in euglenids, the sister group to kinetoplastids, would shed light on the origin of the latter groups's bizarre mtDNA structure and RNA processing. Therefore, it was quite surprising when the mitochondrial genome of *E. gracilis* was recently shown to be extremely streamlined, without any evidence of RNA editing (Dobáková et al. 2015) (Fig. 6.5). This mitochondrial genome consists of a heterogeneous population of 1 to 9 kbp-long linear fragments. Up to now, only seven protein-coding genes have been discovered, as well as two mito-rRNAs (mtSSU and mtLSU), which are each split into two fragments (Spencer and Gray 2011; Dobáková et al. 2015) (Fig. 6.4).

Nonetheless, transmission electron microscopy of the early-branching euglenid *P. cantuscygni* revealed a structure in its mitochondrion resembling the kinetoplast of the kinetoplastid flagellates (Leander et al. 2001; Lee and Simpson 2014). Observations of the mtDNA fraction by electron microscopy confirmed that linear DNA molecules are most frequent, but also small (1 to 2.5 kbp) and large (~40 kbp) circular molecules have been infrequently noted. This observation together with the absence in the sequenced mtDNA segments of some highly conserved mitochondrion-encoded subunits of respiratory complexes III and IV suggest that some kind of RNA editing and gene encryption may exist in this species (Roy et al. 2007).

6.5 Genetic Manipulation of *D. papillatum*

The recently recognized diversity and abundance of diplonemids (Flegontova et al. 2016; Gawryluk et al. 2016) makes it mandatory to turn at least one species into a genetically tractable organism. Indeed, in order to understand their biology, interactions, ecology, and more specifically functions of individual proteins, a crucial step is to establish protocols that would allow genetic manipulations of diplonemids. We have started to develop a transformation system of the type species *D. papillatum*, the genome of which is being sequenced (our unpublished data). Even more

importantly, it can be easily cultivated axenically in the laboratory, reaches high cell density, grows in large volumes, and can be cryopreserved.

Nuclear gene expression of *D. papillatum* is similar to that in other euglenozoans. Its genes are transcribed polycistronically, and individual mRNAs are then trans-spliced, with the short spliced-leader (SL) RNA gene being added to the 5′ end of each transcript. On one hand, the 39-nt-long SL RNA of *D. papillatum* is quite conserved at the sequence level even in the planktonic diplonemids from the DSPD clade. On the other hand, the situation seems much more complex when it comes to nuclear spliceosomal introns, as the nuclear DNA of the DSPD species displays a high density of noncanonical introns that await further characterization (Gawryluk et al. 2016). The genome and transcriptome of *D. papillatum* have been sequenced, and their assembly and annotation are under way (our unpublished data). Knowing the full set of genes will be essential not only for turning this diplonemid into a model species but also for our understanding of its metabolism and other features.

The first obvious task is to get foreign DNA into the *D. papillatum* cells. To ensure stable integration, several crucial steps have to be fulfilled. One is to find resistance markers that can be used for selection of transformants. In the next step, optimal transformation conditions and strategy have to be designed. Last but not least, constructs have to be obtained that will not only stably integrate into the genome, but even more importantly, allow expression, including transcription, posttranscriptional processing and modifications, so that the ensuing transcripts can be finally translated on cytosolic ribosomes. We have accomplished all these steps (Kaur et al. 2018), although efficiency is still moderate and requires optimization.

More specifically, so far we have found seven selection markers to which *D. papillatum* is sensitive. Using available genomic data, we have selected genes that are suited for replacement, namely, those that are nonessential are highly expressed and contain 5′ and 3′ untranslated regions (UTRs) longer than 100 nucleotides. Moreover, we have established a protocol for DNA uptake in a reproducible fashion and have created linear constructs bearing fluorescent protein and selection marker flanked by diplonemid 5′ and 3′ UTRs. We have also confirmed stable incorporation of foreign DNA into the *D. papillatum* genome and have evidence that both the fluorescence gene and the resistance marker on the electroporated constructs are transcribed. Sequencing results showed that the SL RNA sequence is trans-spliced to the 5′ end of the corresponding transcripts. The antibiotic resistance of selected clones provides indirect evidence that the integrated genes are translated (Kaur et al. 2018).

In principle, homologous recombination should be possible, since the genes involved in the corresponding machinery are present in the *D. papillatum* genome, but so far, the inserted DNA has failed to integrate into the target locus. We believe that this can be remedied by further extension of the 5′ and 3′ homologous regions of the constructs. We also plan to use the CRISPR/Cas9-based approach to achieve proper integration of the introduced genes. Attempts to maintain circular plasmids as non-integrated episomes, or to transform the cells with a virus vector carrying green fluorescent protein, were not successful (our unpublished data).

These preliminary data allow us to state that *D. papillatum* can be transformed and has a solid potential to become a genetically tractable organism. Once a robust, reproducible transfection protocol for gene replacement and tagging has been established in *D. papillatum,* we plan to apply the procedure to other diplonemid species—key to understanding the biology of the group as a whole. For the time being, with a representative of the species-rich DSPD clade yet to be brought into culture, the next candidate for transformation is *H. phaeocysticola*. However, this species is much more challenging to work with, as it prefers live diatoms as a food source and reaches only low cell densities. Moreover, in contrast to *D. papillatum*, *H. phaeocysticola* can apparently not be cryopreserved (our unpublished data).

6.6 Conclusions and Outlook

Within the last couple of years, diplonemids have emerged from obscurity as one of the most diverse groups of marine eukaryotes. They are also among the half dozen most abundant eukaryotes. Since their cell numbers seem to expand with depth, one can expect that the importance of diplonemids for the marine ecosystem is widely underappreciated.

Two steps are key for further exploration of these fascinating and ecologically highly relevant protists: (1) complete genome and transcriptome sequences from a broad range of diplonemid species have to become available, and (2) diplonemid species must become amenable to reverse genetic methods, allowing stable integration, transcription, and translation of introduced genes. Given the steadily growing interest in diplonemids, we are optimistic on both accounts.

Acknowledgments We thank Daria Tashyreva, Galina Prokopchuk, and Anzhelika Butenko (Institute of Parasitology) for sharing unpublished data. Support from the Grant Agency of University of South Bohemia (050/2016/P to BK), the Czech Grant Agency (15-21974S and 16-18699S to JL), the ERC CZ (LL1601 to JL), the Canadian Institutes of Health Research (CIHR, MOP 70309 to GB), and from the Gordon and Betty Moore Foundation (GBMF4983.01 to GB and JL) is kindly acknowledged.

References

Abad MG, Long Y, Willcox A, Gott JM, Gray MW, Jackman JE (2011) A role for tRNA(his) guanylyltransferase (Thg1)-like proteins from *Dictyostelium discoideum* in mitochondrial 5′-tRNA editing. RNA 17:613–623

Adl SM, Simpson AGB, Lane CE, Lukeš J, Bass D, Bowser SS, Brown MW, Burki F, Dunthorn M, Hampl V, Heiss A, Hoppenrath M, Lara E, Le Gall L, Lynn DH, McManus H, Mitchell EA, Mozley-Stanridge SE, Parfrey LW, Pawlowski J, Rueckert S, Shadwick L, Schoch CL, Smirnov A, Spiegel FW (2012) The revised classification of eukaryotes. J Eukaryot Microbiol 59:429–493

Adler BK, Harris ME, Bertrand KI, Hajduk SL (1991) Modification of *Trypanosoma brucei* mitochondrial rRNA by posttranscriptional 3′ polyuridine tail formation. Mol Cell Biol 11:5878–5884

Alfonzo JD, Söll D (2009) Mitochondrial tRNA import—the challenge to understand has just begun. Biol Chem 390:717–722

Alfonzo JD, Thiemann O, Simpson L (1997) The mechanism of U insertion/deletion RNA editing in kinetoplastid mitochondria. Nucleic Acids Res 25:3571–3579

Ammerman ML, Downey KM, Hashimi H, Fisk JC, Tomasello DL, Faktorová D, Kafková L, King T, Lukeš J, Read LR (2012) Architecture of the trypanosome RNA editing accessory complex, MRB1. Nucleic Acids Res 40:5637–5650

Aphasizhev R, Aphasizheva I (2011) Uridine insertion/deletion editing in trypanosomes: a playground for RNA-guided information transfer. WIREs 2:669–685

Ban T, Zhu JK, Melcher K, Xu HE (2015) Structural mechanisms of RNA recognition: sequence-specific and non-specific RNA-binding proteins and the Cas9-RNA-DNA complex. Cell Mol Life Sci 72:1045–1058

Benne R, van den Burg J, Brakenhoff JP, Sloof P, van Boom JH, Tromp MC (1986) Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. Cell 46:819–826

Breglia SA, Yubuki N, Hoppenrath M, Leander BS (2010) Ultrastructure and molecular phylogenetic position of a novel euglenozoan with extrusive episymbiotic bacteria: *Bihospites bacati* n. Gen. Et sp. (Symbiontida). BMC Microbiol 10:145

Bundschuh R, Altmüller J, Becker C, Nürnberg P, Gott JM (2011) Complete characterization of the edited transcriptome of the mitochondrion of *Physarum polycephalum* using deep sequencing of RNA. Nucleic Acids Res 39:6044–6055

Burger G, Gray MW, Lang BF (2003) Mitochondrial genomes: anything goes. Trends Genet 19:709–716

Burger G, Gray MW, Forget L, Lang BF (2013) Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists. Genome Biol Evol 5:418–438

Burger G, Moreira S, Valach M (2016) Genes in hiding. Trends Genet 32:553–565

Calvo SE, Clauser KR, Mootha VK (2016) MitoCarta2.0: an updated inventory of mammalian mitochondrial proteins. Nucleic Acids Res 44:D1251–D1257

Cavalier-Smith T (2016) Higher classification and phylogeny of Euglenozoa. Eur J Protistol 56:250–276

Chaput H, Wang Y, Morse D (2002) Polyadenylated transcripts containing random gene fragments are expressed in dinoflagellate mitochondria. Protist 153:111–122

Chen C, Frankhouser D, Bundschuh R (2012) Comparison of insertional RNA editing in *Myxomycetes*. PLoS Comput Biol 8:e1002400

de Vargas C, Audic S, Henry N, Decelle J, Mahé F, Logares R, Lara E, Berney C, Le Bescot N, Probert I, Carmichael M, Poulain J, Romac S, Colin S, Aury J-M, Bittner L, Chaffron S, Dunthorn M, Engelen S, Flegontova O, Guidi L, Horák A, Jaillon O, Lima Mendez G, Lukeš J, Malviya S, Morard R, Mulot M, Scalco E, Siano R, Vincent F, Zingone A, Dimier C, Picheral M, Searson S, Kandels-Lewis S, Acinas SG, Bork P, Bowler C, Gaill F, Gorsky G, Grimsley N, Hingcamp P, Iudicone D, Not F, Ogata H, Pesant S, Raes J, Sieracki M, Speich S, Stemmann L, Sunagawa S, Weissenbach J, Wincker P, Karsenti E, Boss E, Follows M, Karp-Boss L, Krzic U, Reynaud EG, Sardet C, Sullivan MB, Velayoudon D (2015) Eukaryotic plankton diversity in the sunlit global ocean. Science 348:1261605

Dejung M, Subota I, Bucerius F, Dindar G, Freiwald A, Engstler M, Boshart M, Butter F, Janzen CJ (2016) Quantitative proteomics uncovers novel factors involved in developmental differentiation of *Trypanosoma brucei*. PLoS Pathog 12:e1005439

Desmond E, Brochier-Armanet C, Forterre P, Gribaldo S (2011) On the last common ancestor and early evolution of eukaryotes: reconstructing the history of mitochondrial ribosomes. Res Microbiol 162:53–70

Dixit S, Müller-McNicoll M, David V, Zarnack K, Ule J, Hashimi H, Lukeš J (2017) Differential binding of mitochondrial transcripts by MRB8170 and MRB4160 regulates distinct editing fates of mitochondrial mRNA in trypanosomes. MBio 8:e02288-16

Dobáková E, Flegontov P, Skalický T, Lukeš J (2015) Unexpectedly streamlined mitochondrial genome of the euglenozoan *Euglena gracilis*. Genome Biol Evol 7:3358–3367

Edgcomb VP, Breglia SA, Yubuki N, Beaudoin D, Patterson DJ, Leander BS, Bernhard JM (2011) Identity of epibiotic bacteria on symbiontid euglenozoans in O2-depleted marine sediments: evidence for symbiont and host co-evolution. ISME J 5:231–243

Eloe EA, Shulse CN, Fadrosh DW, Williamson SJ, Allen EE, Bartlett DH (2011) Compositional differences in particle-associated and free-living microbial assemblages from an extreme deep-ocean environment. Environ Microbiol Rep 3:449–458

El-Sayed NM, Myler PJ, Blandin G, Berriman M, Crabtree J, Aggarwal G, Caler E, Renauld H, Worthey EA, Hertz-Fowler C, Ghedin E, Peacock C, Bartholomeu DC, Hass BJ, Tran A-N, Wortman JR, Alsmark UCM, Angiuoli S, Anupama A, Badger J, Bringaud F, Cadag E, Carlton JM, Cerqueira GC, Creasy T, Delcher AL, Djikeng A, Embley TM, Hauser C, Ivens AC, Kummerfeld SK, Pereira-Leal JB, Nilsoon D, Peterson J, Salzberg SL, Shallom J, Silva JC, Sundaram J, Westenberger S, White O, Melville SE, Donelson JE, Andersson B, Stuart KD, Hall N (2005) Comparative genomics of trypanosomatid parasitic protozoa. Science 309:404–409

Faktorová D, Dobáková E, Peña-Diaz P, Lukeš J (2016) From simple to supercomplex: mitochondrial genomes of euglenozoan protists. F1000 Res 5:392

Flegontov P, Gray MW, Burger G, Lukeš J (2011) Gene fragmentation: a key to mitochondrial genome evolution in Euglenozoa? Curr Genet 57:225–232

Flegontov P, Michálek J, Janouškovec J, Lai D-H, Jirků M, Hajdušková E, Tomčala A, Otto TD, Keeling PJ, Pain A, Oborník M, Lukeš J (2015) Divergent mitochondrial respiratory chains in phototrophic relatives of apicomplexan parasites. Mol Biol Evol 32:1115–1131

Flegontova O, Flegontov P, Malviya S, Audic S, Wincker P, de Vargas C, Bowler C, Lukeš J, Horák A (2016) Unexpected diversity and abundance of planktonic diplonemids in the world ocean. Curr Biol 26:3060–3065

Fu C-J, Sheikh S, Miao W, Andersson SGE, Baldauf SL (2014) Missing genes, multiple ORFs, and C-to-U type RNA editing in *Acrasis kona* (Heterolobosea, Discoba) mitochondrial DNA. Genome Biol Evol 6:2240–2257

Gawryluk RMR, Chisholm KA, Pinto DM, Gray MW (2014) Compositional complexity of the mitochondrial proteome of a unicellular eukaryote (*Acanthamoeba castellanii*, supergroup Amoebozoa) rivals that of animals, fungi, and plants. J Proteome 109:400–416

Gawryluk RMR, del Campo J, Okamoto N, Strassert JFH, Lukeš J, Richards TA, Worden AZ, Santoro AE, Keeling PJ (2016) Morphological identification and single-cell genomics of marine diplonemids. Curr Biol 26:3053–3059

Gonzalez TN, Sidrauski C, Dörfler S, Walter P (1999) Mechanism of non-spliceosomal mRNA splicing in the unfolded protein response pathway. EMBO J 18:3119–3132

Göringer HU (2012) 'Gestalt,' composition and function of the *Trypanosoma brucei* editosome. Annu Rev Microbiol 66:65–82

Gott JM, Parimi N, Bundschuh R (2005) Discovery of new genes and deletion editing in *Physarum* mitochondria enabled by a novel algorithm for finding edited mRNAs. Nucleic Acids Res 33:5063–5072

Gray MW (2012) Mitochondrial evolution. Cold Spring Harb Perspect Biol 4:a011403

Hudson AJ, Stark MR, Fast NM, Russell AG, Rader SD (2015) Splicing diversity revealed by reduced spliceosomes in *C. merolae* and other organisms. RNA Biol 12:1–8

Jackman JE, Gott JM, Gray MW (2012) Doing it in reverse: 3′-to-5′ polymerization by the Thg1 superfamily. RNA 18:886–899

Jackson CJ, Waller RF (2013) A widespread and unusual RNA trans-splicing type in dinoflagellate mitochondria. PLoS One 8:e56777

Jackson CJ, Norman JE, Schnare MN, Gray MW, Keeling PJ, Waller RF (2007) Broad genomic and transcriptional analysis reveals a highly derived genome in dinoflagellate mitochondria. BMC Biol 5:41

Karnkowska A, Vacek V, Zubáčová Z, Treitli SC, Petrželková R, Eme L, Novák L, Žárský V, Barlow LD, Herman EK, Soukal P, Hroudová M, Doležal P, Stairs CW, Roger AJ, Eliáš M, Dacks JB, Vlček Č, Hampl V (2016) A eukaryote without a mitochondrial organelle. Curr Biol 26:1274–1284

Kaur B, Valach M, Peña-Diaz P, Moreira S, Keeling PJ, Burger G, Lukeš J, Faktorová D (2018) Transformation of *Diplonema papillatum*, the type species of the highly diverse and abundant marine microeukaryotes Diplonemida (Euglenozoa). Environ Microbiol 20:1030–1040

Kiethega GN, Turcotte M, Burger G (2011) Evolutionary conserved *cox1* trans-splicing without cis-motifs. Mol Biol Evol 28:2425–2458

Kiethega GN, Yan Y, Turcotte M, Burger G (2013) RNA-level unscrambling of fragmented genes in *Diplonema* mitochondria. RNA Biol 10:301–313

Knoop V (2010) When you can't trust the DNA: RNA editing changes transcript sequences. Cell Mol Life Sci 68:567–586

Kuai L, Fang F, Butler JS, Sherman F (2004) Polyadenylation of rRNA in *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A 101:8581–8586

Laforest MJ, Roewer I, Lang BF (1997) Mitochondrial tRNAs in the lower fungus *Spizellomyces punctatus*: tRNA editing and UAG 'stop' codons recognized as leucine. Nucleic Acids Res 25:626–632

Lara E, Moreira D, Vereshchaka A, López-García P (2009) Panoceanic distribution of new highly diverse clades of deep-sea diplonemids. Environ Microbiol 11:47–55

Lavrov DV, Adamski M, Chevaldonné P, Adamska M (2016) Extensive mitochondrial mRNA editing and unusual mitochondrial genome organization in calcaronean sponges. Curr Biol 26:86–92

Leander BS, Triemer RE, Farmer MA (2001) Character evolution in heterotrophic euglenids. Eur J Protistol 37:337–356

Lee WJ, Simpson AGB (2014) Morphological and molecular characterisation of *Notosolenus urceolatus* Larsen and Patterson 1990, a member of an understudied deepbranching euglenid group (petalomonads). J Eukaryot Microbiol 61:463–479

Leigh J, Lang BF (2004) Mitochondrial 3′ tRNA editing in the jakobid *Seculamonas ecuadoriensis*: a novel mechanism and implications for tRNA processing. RNA 10:615–621

Levin JZ, Yassour M, Adiconis X, Nusbaum C, Thompson DA, Friedman N, Gnirke A, Regev A (2010) Comprehensive comparative analysis of strand-specific RNA sequencing methods. Nat Methods 7:709–715

Lin S, Zhang H, Spencer DF, Norman JE, Gray MW (2002) Widespread and extensive editing of mitochondrial mRNAs in dinoflagellates. J Mol Biol 320:727–739

Lithgow T, Schneider A (2010) Evolution of macromolecular import pathways in mitochondria, hydrogenosomes and mitosomes. Philos Trans R Soc B Sci 365:799–817

López-García P, Rodríguez-Valera F, Pedrós-Alió C, Moreira D (2001) Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. Nature 409:603–607

López-García P, Vereshchaka A, Moreira D (2007) Eukaryotic diversity associated with carbonates and fluid seawater interface in Lost City hydrothermal field. Environ Microbiol 9:546–554

Lukeš J, Archibald JM, Keeling PJ, Doolittle WF, Gray MW (2011) How a neutral evolutionary ratchet can build cellular complexity. IUBMB Life 63:528–537

Lukeš J, Flegontova O, Horák A (2015) Diplonemids. Curr Biol 25:R702–R704

Maguire F, Richards TA (2014) Organelle evolution: a mosaic of 'mitochondrial' functions. Curr Biol 24:R518–R520

Mahendran R, Spottswood MR, Miller DL (1991) RNA editing by cytidine insertion in mitochondria of *Physarum polycephalum*. Nature 349:434–438

Manna S (2015) An overview of pentatricopeptide repeat proteins and their applications. Biochimie 113:93–99

Marande W, Burger G (2007) Mitochondrial DNA as a genomic jigsaw puzzle. Science 318:415
Marande W, Lukeš J, Burger G (2005) Unique mitochondrial genome structure in diplonemids, the sister group of kinetoplastids. Eukaryot Cell 4:1137–1146
Maslov DA, Yasuhira S, Simpson L (1999) Phylogenetic affinities of *Diplonema* within the Euglenozoa as inferred from the SSU rRNA gene and partial COI protein sequences. Protist 150:33–42
Mohanty BK, Kushner SR (2011) Bacterial/archaeal/organellar polyadenylation. WIREs 2:256–276
Moreira S, Breton S, Burger G (2012) Unscrambling of genetic information at the RNA level. WIREs 3:213–228
Moreira S, Valach M, Aoulad-Aissa M, Otto C, Burger G (2016) Novel modes of RNA editing in mitochondria. Nucleic Acids Res 44:4907–4919
Nishikura K (2016) A-to-I editing of coding and non-coding RNAs by ADARs. Nat Rev Mol Cell Biol 17:83–96
Okamoto N, Gawryluk RMR, del Campo J, Strassert JFH, Lukeš J, Richards TA, Worden AZ, Santoro AE, Keeling PJ (2018) *Eupelagonema oceanica* n. gen. & sp. and a revised diplonemid taxonomy. J Eukaryot Microbiol (in press)
Ozsolak F, Milos PM (2011) RNA sequencing: advances, challenges and opportunities. Nat Rev Genet 12:87–98
Parkhomchuk D, Borodina T, Amstislavskiy V, Banaru M, Hallen L, Krobitsch S, Lehrach H, Soldatov A (2009) Transcriptome analysis by strand-specific sequencing of complementary DNA. Nucleic Acids Res 37:e123
Pawlowski J, Audic S, Adl S, Bass D, Belbahri L, Berney C, Bowser SS, Čepička I, Decelle J, Dunthorn M, Fiore-Donno A-M, Gile HG, Holzmann M, Jahn R, Jirků M, Keeling PJ, Kostka M, Kudryavtsev A, Lara E, Lukeš J, Mann GD, Mitchell ADE, Nitsche F, Romeralo M, Saunders WG, Simpson AGB, Smirnov VA, Spouge J, Stern FR, Stoeck T, Zimmermann J, Schindel D, de Vargas C (2012) CBOL Protist working group: barcoding eukaryotic richness beyond the animal, plant and fungal kingdoms. PLoS Biol 10:e1001419
Popow J, Schleiffer A, Martinez J (2012) Diversity and roles of (t)RNA ligases. Cell Mol Life Sci 69:2657–2670
Rammelt C, Rossmanith W (2016) Repairing tRNA termini: news from the 3′ end. RNA Biol 13:1182–1188
Read LK, Lukeš J, Hashimi H (2016) Trypanosome RNA editing: the complexity of getting U in and taking U out. WIREs 7:33–51
Roy J, Faktorová D, Lukeš J, Burger G (2007) Unusual mitochondrial genome structures throughout the Euglenozoa. Protist 158:385–396
Rubio MAT, Pastar I, Gaston KW, Ragone FL, Janzen CJ, Cross GAM, Papavasiliou FN, Alfonzo JD (2007) An adenosine-to-inosine tRNA-editing enzyme that can perform C-to-U deamination of DNA. Proc Natl Acad Sci U S A 104:7821–7826
Rüdinger M, Fritz-Laylin L, Polsakiewicz M, Knoop V (2011) Plant-type mitochondrial RNA editing in the protist *Naegleria gruberi*. RNA 17:2058–2062
Salone V, Rüdinger M, Polsakiewicz M, Hoffmann B, Groth-Malonek M, Szurek B, Small I, Knoop V, Lurin C (2007) A hypothesis on the identification of the editing enzyme in plant organelles. FEBS Lett 581:4132–4138
Schnepf E (1994) Light and electron microscopical observations in *Rhynchopus coscinodiscivorus* spec. Nov., a colorless, phagotrophic Euglenozoon with concealed flagella. Arch Protistenkd 144:63–74
Segovia R, Pett W, Trewick S, Lavrov DV (2011) Extensive and evolutionarily persistent mitochondrial tRNA editing in velvet worms (phylum Onychophora). Mol Biol Evol 28:2873–2881
Shapiro TA, Englund PT (1995) The structure and replication of kinetoplast DNA. Annu Rev Microbiol 49:117–143
Shikanai T (2015) RNA editing in plants: machinery and flexibility of site recognition. Biochim Biophys Acta 1847:779–785

Škodová-Sveráková I, Verner Z, Skalický T, Votýpka J, Horváth A, Lukeš J (2015) Lineage-specific activities of a multipotent mitochondrion of trypanosomatid flagellates. Mol Microbiol 96:55–67

Slomovic S, Fremder E, Staals RH, Pruijn GJ, Schuster G (2010) Addition of poly(a) and poly(a)-rich tails during RNA degradation in the cytoplasm of human cells. Proc Natl Acad Sci U S A 107:7407–7412

Smith DR, Keeling PJ (2015) Mitochondrial and plastid genome architecture: reoccurring themes, but significant differences at the extremes. Proc Natl Acad Sci U S A 112:10177–10184

Spencer DF, Gray MW (2011) Ribosomal RNA genes in *Euglena gracilis* mitochondrial DNA: fragmented genes in a seemingly fragmented genome. Mol Gen Genomics 285:19–31

Stahley MR, Strobel SA (2006) RNA splicing: group I intron crystal structures reveal the basis of splice site selection and metal ion catalysis. Curr Opin Struct Biol 16:319–326

Stuart K, Feagin JE (1992) Mitochondrial DNA of kinetoplastids. Int Rev Cytol 141:65–88

Sun T, Bentolila S, Hanson MR (2016) The unexpected diversity of plant organelle RNA editosomes. Trends Plant Sci 21:962–973

Szklarczyk R, Huynen MA (2010) Mosaic origin of the mitochondrial proteome. Proteomics 10:4012–4024

Takenaka M, Zehrmann A, Verbitskiy D, Härtel B, Brennicke A (2013) RNA editing in plants and its evolution. Annu Rev Genet 47:335–352

Tanaka N, Meineke B, Shuman S (2011) RtcB, a novel RNA ligase, can catalyze tRNA splicing and HAC1 mRNA splicing *in vivo*. J Biol Chem 286:30253–30257

Tashyreva D, Prokopchuk G, Yabuki A, Kaur B, Faktorová D, Votýpka J, Kusaka C, Fujikura K, Shiratori T, Ishida K, Horák A, Lukeš J (2018) Phylogeny and morphology of diplonemids from the Sea of Japan. Protist 169:158–179

Tielens AGM, van Hellemond JJ (2009) Surprising variety in energy metabolism within Trypanosomatidae. Trends Parasitol 25:482–490

Triemer RE, Ott DW (1990) Ultrastructure of *Diplonema ambulator* Larsen & Patterson (Euglenozoa) and its relationship to *Isonema*. Eur J Protistol 25:316–320

Urbaniak MD, Martin DM, Ferguson MA (2013) Global quantitative SILAC phosphoproteomics reveals differential phosphorylation is widespread between the procyclic and bloodstream form lifecycle stages of *Trypanosoma brucei*. J Proteome Res 12:2233–2244

Valach M, Moreira S, Kiethega GN, Burger G (2014) Trans splicing and RNA editing of LSU rRNA in *Diplonema* mitochondria. Nucleic Acids Res 42:2660–2672

Valach M, Moreira S, Faktorová D, Lukeš J, Burger G (2016) Post-transcriptional mending of gene sequences: looking under the hood of mitochondrial gene expression in diplonemids. RNA Biol 13:1204–1211

Valach M, Moreira S, Hoffmann S, Stadler PF, Burger G (2017) Keeping it complicated: mitochondrial genome plasticity in diplonemids. Sci Rep 7:14166

Vanfleteren JR, Vierstraete AR (1999) Insertional RNA editing in metazoan mitochondria: the cytochrome b gene in the nematode *Teratocephalus lirellus*. RNA 5:622–624

Verner Z, Basu S, Benz C, Dixit S, Dobáková E, Faktorová D, Hashimi H, Horáková E, Huang Z, Paris Z, Pena-Diaz P, Ridlon L, Týč J, Wildridge D, Zíková A, Lukeš J (2015) Malleable mitochondrion of *Trypanosoma brucei*. Int Rev Cell Mol Biol 315:73–151

Visomirski-Robic LM, Gott JM (1997) Insertional editing of nascent mitochondrial RNAs in *Physarum*. Proc Natl Acad Sci U S A 94:4324–4329

Vlcek Č, Marande W, Teijeiro S, Lukeš J, Burger G (2011) Gene fragments scattered across a multipartite mitochondrial genome. Nucleic Acids Res 39:979–988

Wahlstedt H, Ohman M (2011) Site-selective versus promiscuous A-to-I editing. WIREs 2:761–771

Wheeler RJ, Gull K, Gluenz E (2012) Detailed interrogation of trypanosome cell biology via differential organelle staining and automated image analysis. BMC Biol 10:1

Yabuki A, Tame A (2015) Phylogeny and reclassification of *Hemistasia phaeocysticola* (Scherffel) Elbrachter & Schnepf, 1996. J Eukaryot Microbiol 62:426–429

Yabuki A, Tanifuji G, Kusaka C, Takishita K, Fujikura K (2016) Hyper-eccentric structural genes in the mitochondrial genome of the algal parasite *Hemistasia phaeocysticola*. Genome Biol Evol:8, 2870–2878

Yang J, Harding T, Kamikawa R, Simpson AGB, Roger AJ (2017) Mitochondrial genome evolution and a novel RNA editing system in deep-branching heteroloboseids. Genome Biol Evol 9:1161–1174

Yokobori S, Pääbo S (1995) Transfer RNA editing in land snail mitochondria. Proc Natl Acad Sci U S A 92:10432–10435

Yubuki N, Edgcomb VP, Bernhard JM, Leander BS (2009) Ultrastructure and molecular phylogeny of *Calkinsia aureus*: cellular identity of a novel clade of deep-sea euglenozoans with epibiotic bacteria. BMC Microbiol 27:16

Yubuki N, Simpson AG, Leander BS (2013) Reconstruction of the feeding apparatus in *Postgaardi mariagerensis* provides evidence for character evolution within the Symbiontida (Euglenozoa). Eur J Protistol 49:32–39

Zhao C, Pyle AM (2017) Structural insights into the mechanism of group II intron splicing. Trends Biochem Sci 42:470–482

Zíková A, Verner Z, Nenarokova A, Michels PAM, Lukeš J (2017) A paradigm shift: the mitoproteomes of procyclic and bloodstream *Trypanosoma brucei* are comparably complex. PLoS Pathog 13:e1006679

Zimorski V, Ku C, Martin W, Gould SB (2014) Endosymbiotic theory for organelle origins. Curr Opin Microbiol 22:38–48

Chapter 7
Mechanisms and Evolution of tRNA 5′-Editing in Mitochondria

Samantha Dodbele, Jane E. Jackman, and Michael W. Gray

Abstract In several protists and fungi, many of the tRNAs encoded by the mitochondrial genome are unusual in that they are predicted to have mismatches within the first three positions of the acceptor stem. However, examination of the sequences of the corresponding mature tRNAs has shown that these positions instead contain canonical Watson-Crick-type base pairs. This difference results from changes that are made at the transcript level, such that predicted mismatches are effectively corrected. The correction process, termed mitochondrial tRNA 5′-editing (*mt-tRNA 5′-editing*), involves removal in the 5′-to-3′ direction of several nucleotides, starting at the 5′-end of the acceptor stem and including those 5′ nucleotides at positions of mismatching, followed by sequential addition of nucleotides in the 3′-to-5′ direction to fill in the resulting gap, with nucleotides on the 3′ side of the stem serving to guide incorporation. While the nature of the nuclease(s) involved in removal of nucleotides during mt-tRNA 5′-editing is unknown, the addition function is carried out by a mitochondrion-targeted Thg1-like protein (TLP), a novel 3′-to-5′ nucleotidyltransferase ("reverse RNA polymerase"). Thg1 (tRNA-histidine guanylyltransferase), the founding member of the protein family to which TLPs also belong, catalyzes the addition of a single, non-templated G residue to the 5′-end of histidine tRNA, whereas TLPs involved in mt-tRNA 5′-editing robustly catalyze multiple rounds of templated addition of nucleotides to the 5′-end of appropriately truncated tRNA substrates. To date, mt-tRNA 5′-editing has been experimentally documented in several amoebozoan and fungal species and is predicted to occur in several other protist lineages. Consideration of phylogenetic distribution and biochemical characteristics suggests a constructive neutral evolution (CNE) scenario for the evolution of mt-tRNA 5′-editing, wherein mitochondrion-targeted TLPs independently

S. Dodbele · J. E. Jackman
Department of Chemistry and Biochemistry, The Ohio State Biochemistry Program, Center for RNA Biology, The Ohio State University, Columbus, OH, USA

M. W. Gray (✉)
Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, NS, Canada

Center for Comparative Genomics and Evolutionary Bioinformatics, Dalhousie University, Halifax, NS, Canada
e-mail: m.w.gray@dal.ca

J. Cruz-Reyes, M. Gray (eds.), *RNA Metabolism in Mitochondria*, Nucleic Acids and Molecular Biology 34, https://doi.org/10.1007/978-3-319-78190-7_7

emerge in discrete eukaryotic lineages, thereby allowing the fixation in the mitochondrial genome of tRNA mismatch mutations that would otherwise be purged by purifying selection.

7.1 Introduction

The term *RNA editing* encompasses a wide variety of mechanistically and evolutionarily unrelated processes that re-tailor RNA transcripts so that the final nucleotide sequence differs from that of the corresponding DNA coding element (Gray 2003). Two distinct categories of editing are recognized (*substitution* and *insertion-deletion*), depending on whether or not, respectively, the sequence of the final edited RNA is co-linear, nucleotide-for-nucleotide, with that of the encoding DNA. (Note that this categorization only addresses the state of the final edited product but says nothing about the actual mechanism of editing: both substitution and insertion-deletion types of editing may proceed by quite different biochemical pathways in the different organisms in which they occur.) In type and extent, RNA editing is almost exclusively a eukaryotic phenomenon and is particularly prominent in eukaryotic organelles, plastids, and mitochondria, especially the latter (Gray 2012).

Messenger RNA (mRNA) was the first class of RNA for which editing was discovered (Benne et al. 1986; Chen et al. 1987; Powell et al. 1987). Subsequently, other classes of RNA (ribosomal, transfer, viral) were shown to undergo editing, with the first reports of transfer RNA (tRNA) editing appearing in 1993 (see Price and Gray 1998).

Editing of tRNAs in the mitochondria of *Acanthamoeba castellanii* (Lonergan and Gray 1993a, b) provided the first example of what has come to be known as "mt-tRNA 5′-editing." This type of editing was discovered during sequencing of the mitochondrial genome of *A. castellanii*, a single-celled eukaryotic microbe (protist). *A. castellanii* mitochondrial DNA (mtDNA) encodes 15 distinct tRNA species (Burger et al. 1995), and secondary structure modeling predicted that 12 of these tRNAs would have one or more mismatches in the first three positions of the acceptor stem, normally a helix consisting of 7 base pairs. This unprecedented situation raised the question of whether mt-tRNAs having a mismatched acceptor stem could actually be functional, considering the essential role of this helical element in forming the L-shaped three-dimensional tRNA structure (Kim 1978).

However, sequencing of the corresponding mature tRNAs demonstrated that these predicted mismatches had been corrected posttranscriptionally during tRNA maturation by changes (editing) within the 5′ half of the acceptor stem, generating standard Watson-Crick (WC) base pairs (Lonergan and Gray 1993a). In all, 21 edits were predicted by secondary structure modeling, all of which were subsequently confirmed by sequencing the corresponding mature tRNAs (Lonergan and Gray 1993a; Price and Gray 1999a). Unexpectedly, two U∘G pairs (at position 1 in a $tRNA^{Met}$ and position 3 in a $tRNA^{Leu2}$) were also seen to be edited to standard C-G pairs, even though both U∘G and G∘U pairs are tolerated within and frequently

found in RNA helices. U∘G/G∘U editing only appeared to occur within the first three positions of the acceptor stem: pairs of this type elsewhere within the acceptor stem were unaffected.

From these initial studies, several properties of this editing system emerged. (1) Mismatches as well as noncanonical U∘G and G∘U pairs within the first three positions of the acceptor stem (but not elsewhere) are converted to standard WC base pairs via appropriate nucleotide substitutions on the 5′ side of the stem. (2) This pattern implies that the 3′ half of the acceptor stem effectively serves as an internal guide for editing on the 5′ side of the stem. (3) Pyrimidine-to-purine, purine-to-purine, and pyrimidine-to-purine edits occur, in a ratio of 15:6:2 in this particular mitochondrial system. (4) The overwhelming number of changes are from U, A, or C to G: i.e., in an edited tRNA, the 5′ half of the stem becomes decidedly more G-rich after editing.

Consideration of the substitutions generated by this editing system suggested a mechanism whereby the first three nucleotides on the 5′ half of the acceptor stem, whether mismatched or not, are excised by a nuclease activity (exo- or endo-) and replaced in a stepwise fashion by a 3′-to-5′ nucleotidyltransferase activity, using the corresponding nucleotides on the 3′ half of the stem as template and guide (Lonergan and Gray 1993a): the latter an admittedly unprecedented type of enzyme (Fig. 7.1). Development of an in vitro tRNA editing system starting with purified *A. castellanii*

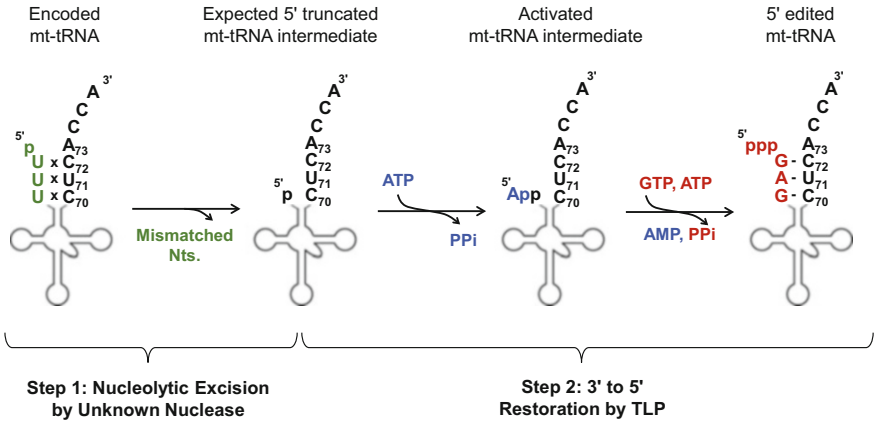

Fig. 7.1 Mitochondrial tRNA 5′-editing pathway. The 5′-editing pathway removes 5′-end mismatched nucleotides (indicated in green in this representative example) by an as-yet unidentified nuclease step. Intermediates that lack 5′-nucleotides such as the expected intermediate shown here have been observed in several sequencing experiments, although other possible intermediate species, such as those with additional correct 5′-nucleotides removed along with the mismatches, cannot be ruled out. The 5′-end repair activity requires first, activation of the 5′-phosphate that results from 5′-mismatch removal, followed by addition of the correct Watson-Crick base pairing nucleotides (GTP and ATP in this example) to restore a fully base-paired 5′-end to the mt-tRNA. In *Dictyostelium discoideum*, this activity is catalyzed by one of the Thg1/TLP family enzymes (DdiTLP3) encoded in this organism. Homologues of DdiTLP3 are likely to catalyze similar roles in other species, although this inference has yet to be experimentally validated

mitochondria provided strong and direct evidence for the existence of such an enzyme (Price and Gray 1999b). This in vitro system was employed in protein fractionation experiments aimed at isolating a purified activity that could be studied further, but these attempts proved unsuccessful, a result attributed in part to an apparently very low concentration of the putative nucleotidyltransferase in active mitochondrial extracts (A. Lohan and M.W. Gray unpublished).

At the time, the only known enzyme acting as a 3′-to-5′-nucleotidyltransferase was histidine tRNA guanylyltransferase, which was named Thg1 upon its discovery in yeast (*Saccharomyces cerevisiae*) nearly two decades later (Gu et al. 2003). This enzyme adds a G residue specifically to the 5′-end of $tRNA^{His}$ opposite the discriminator nucleotide at position N_{73}, the fourth position from the mature 3′ terminus (Cooley et al. 1982; Williams et al. 1990). In eukaryotes, G is added opposite a universally conserved A in the discriminator position. Although the catalytic mechanism of the yeast enzyme had been investigated in partially purified extracts (Jahn and Pande 1991), no sequence information for the corresponding protein was available; hence, no bioinformatic survey could be undertaken to look for potential homologues that might have 3′-to-5′-nucleotidyltransferase activity.

Subsequent identification and sequence determination of the yeast gene for Thg1 revealed that a highly conserved ortholog is encoded by the vast majority of sequenced eukaryotic genomes (Gu et al. 2003; Jackman et al. 2012). Thg1 proved to be the founding member of the Thg1 superfamily (Jackman et al. 2012), whose members include related Thg1-like proteins, designated "TLPs," identified in selected groups of archaea, bacteria, and eukaryotes. As detailed below, TLPs proved to have a robust 3′-to-5′-nucleotidyltransferase activity, able to progressively add nucleotides to the 5′-end of appropriately truncated tRNA substrates.

7.2 Mitochondrial tRNA Editing in *Dictyostelium discoideum*

Dictyostelium discoideum is a soil-dwelling social amoeba commonly utilized as a model organism for motility, chemotaxis, cytokinesis, and phagocytosis studies. Its unique life cycle, transitioning from single cells to a multicellular unit upon starvation, makes it an ideal model for studying cell differentiation and signaling (Ogawa et al. 2000; Eichinger et al. 2005). Shortly after mt-tRNA 5′-editing was discovered in *A. castellanii,* the same type of mt-tRNA editing in *D. discoideum* was inferred after the sequencing of its mitochondrial genome (Ogawa et al. 2000). The mtDNA of *D. discoideum* encodes 18 tRNA genes, with 8 of these predicting mismatches in the first three positions of the acceptor stem, similar to *A. castellanii* (Ogawa et al. 2000; Laforest et al. 2004; Abad et al. 2014). Furthermore mitochondrial genomes of both organisms exhibit similarities suggesting a common ancestral origin, such as similar genome sizes (55.5 kb in *D. discoideum* and 41.6 kb in *A. castellanii*), A + T content, (72.6% in *D. discoideum* and 70.6% in *A. castellanii*), a fused *cox*1/2 gene,

and gene order of the ribosomal protein cluster (Ogawa et al. 2000). Nonetheless, the observation of tRNA mismatches stochastically distributed among different types and positions of the mt-tRNAs suggests that specific changes observed in different mt-tRNA genes are the products of independent evolutionary events (Ogawa et al. 2000; see Sect. 7.5, below).

Initially the eight mismatch-containing mt-tRNA species were predicted to be edited based on the presence of A × C, U × C, A × G, C × A, and A × A encoded mismatches in the acceptor stem sequence (Laforest et al. 2004; Abad et al. 2011, 2014). Editing of these tRNAs was later confirmed through the well-established approach of mt-tRNA circularization followed by RT-PCR (Lohan and Gray 2007; Abad et al. 2014). The resulting sequences revealed that all eight mismatch-containing mt-tRNAs were completely edited with their 5′-stem restored to the correct WC base pairing (Table 7.1). Interestingly, some cDNA sequences derived from two of these mt-tRNAs (Ile_{CAU} and Glu_{UUC}) either retained the mismatch or had the mismatch removed but the correct WC base pairing not yet added (Abad et al. 2014). These partially edited species likely represent intermediates in the editing process and support the originally proposed mechanism whereby one or more nucleases act(s) on the 5′-end of the mt-tRNA to generate a 5′-truncated species that is the substrate for 3′-to-5′ addition by the repair nucleotidyltransferase.

Seven *D. discoideum* mt-tRNAs contain a wobble G∘U/U∘G at various positions within the first three base pairs of the acceptor stem; however, since wobble base pairs are generally tolerated in tRNA, and only Ile_{GAU1} also contains other non-WC mismatches along with the U∘G base pair, this was the only one of these seven mt-tRNAs that was initially expected to undergo 5′-editing. Indeed, editing of this mt-tRNA (Ile_{GAU1}) was confirmed upon sequencing (Abad et al. 2014). However, two additional mt-tRNAs were also revealed to be 5′-editing substrates (Pro_{UGG}, and Trp_{CCA}), with the wobble U∘G/G∘U base pairs in these mt-tRNAs being edited to WC C-G or A-U pairs, respectively (Abad et al. 2014). The difference here appears to be dependent on the location of the wobble base pair, since both of these edited mt-tRNAs contain a G∘U at the first position of the acceptor stem, while the remaining four unedited mt-tRNAs contain the wobble pair at the third position, in the absence of any bona fide mismatches. Nonetheless, this evidence of 5′-editing despite the absence of a traditional mismatch is of particular interest as it suggests that the nuclease component of the 5′-editing machinery does not solely rely upon recognizing specific types of base pairs but in addition uses other aspects of sequence context to effect recognition of "incorrect" nucleotides to be removed (Abad et al. 2014).

Although wobble base pair editing is clearly observed in *D. discoideum* and *A. castellanii* (Price and Gray 1999a; Abad et al. 2014), it is only minimally observed in *Polysphondylium pallidum*, *Monoblepharella*, and *Harpochytrium*, organisms that otherwise contain very similar mismatch and editing patterns in mt-tRNAs, suggesting that 5′-editing may be governed by organism-specific rules (Laforest et al. 2004; Abad et al. 2014; Long and Jackman 2015). A comparison of *D. discoideum* and *P. pallidum* mt-tRNAs that encode G∘U/U∘G base pairs reveals some striking examples of these types of organism-specific patterns of editing. In

Table 7.1 Verified mt-tRNA 5′-editing events in *Amoebozoa*

Organism[a]	mt-tRNA	Position	Edit	Resulting pair[b]	Tested by[c]	Citation
Acanthamoeba castellanii (12/15)	Ile2(CAU)	2	A → G	G–C	RT	Lonergan and Gray (1993a)
		3	U → G	G–C		
	Ala(UGC)	3	U → A	A–U	RT, Cir	Lonergan and Gray (1993a), Price and Gray (1999a)
	Asp(GUC)	1	U → G	G–C	RT, Cir	
		2	U → A	A–U		
		3	U → G	G–C		
	Met(CAU)	1	U → C	C–G	RT, Cir	
		2	U → A	A–U		
	Lys(CUU)	1	A → G	G–C	Cir	Price and Gray (1999a)
		2	U → G	G–C		
		3	U → G	G–C		
	Glu(UUC)	1	A → G	G–C	Cir	
	Ile1(GAU)	1	A → G	G–C	Cir	
		2	C → A	A–U		
	Leu1(UAG)	1	A → G	G–C	Cir	
		2	U → G	G–C		
	Phe(GAA)	1	U → G	G–C	Cir	
		2	C → G	G–C		
		3	A → G	G–C		
	Tyr(GUA)	2	U → G	G–C	Cir	
	Trp(CCA)	3	C → G	G–C	Cir	
	Leu2(UAA)	1	U → G	G–C	Cir	Price and Gray (1999a), Lohan and Gray (2007)
		3	U → C	C–G		
Dictyostelium discoideum (10/18)	Ile(CAU)	1	A → G	G–C	Cir, PPA	Abad et al. (2011, 2014)
	Glu(UUC)	1	U → G	G–C	Cir	Abad et al. (2014)
	Tyr(GUA)	1	A → G	G–C	Cir	
	Leu(UAG)	1	A → G	G–C	Cir, PPA	Abad et al. (2011, 2014)
		2	A → C	C–G		
	Gln(UUG)	1	C → U	U–A	Cir	Abad et al. (2014)
		2	A → C	C–G		
	Ile(GAU2)	1	A → C	C–G	Cir	
		2	A → U	U–A		
	Leu(UAA)	2	A → U	U–A	Cir	
		3	A → C	C–G		
	Ile(GAU1)	1	A → C	C–G	Cir	
		2	A → U	U–A		
		3	U → C	C–G		
	Pro(UGG)	1	U → C	C–G	Cir	
	Trp(CCA)	1	G → A	A–U	Cir	

(continued)

Table 7.1 (continued)

Organism[a]	mt-tRNA	Position	Edit	Resulting pair[b]	Tested by[c]	Citation
Polysphondylium pallidum (5/11)	Ala(UGC)	1	U → G	G–C	Cir	Abad et al. (2014)
	Asn(GUU)	1	A → G	G–C	Cir	
	Ile(GAU)	2	A → C	C–G	Cir	
	Leu(UAA)	1	A → C	C–G	Cir	
		3	A → C	C–G		
	PheGAA	1	A → G	G–C	Cir	
		2	A → U	U–A		
		3	U → C	C–G		
Physarum polycephalum (2/5)	Met1(CAU)	1	U → G	G–C	Cir	Gott et al. (2010)
	Met2(CAU)	1	C → G	G–C	Cir	

[a]Numbers in parentheses indicate no. of edited mt-tRNAs vs. total no. of mt-tRNAs encoded by the mitochondrial genome
[b]Positions 1, 2, and 3 of the acceptor stem comprise paired nucleotides 1–72, 2–71, and 3–70, respectively, according to the standard tRNA numbering
[c]*RT* reverse transcriptase sequencing, *Cir* circularization, *PPA* phosphatase protection assay

D. discoideum mt-tRNA$^{Ile}_{GAU1}$, two canonical mismatches occur immediately adjacent to a U∘G base pair, and all bases including the U∘G base pair are 100% edited (Abad et al. 2014; Long and Jackman 2015). However in *P. pallidum* mt-tRNA$^{Phe}_{GAA}$, where the U∘G base pair occurs at the same location and in a similar context, immediately adjacent to a canonical mismatch, the U∘G wobble pair is nearly 100% unedited (Long and Jackman 2015).

In summary, in *D. discoideum* 10 mt-tRNAs undergo 5′-editing (Leu_{UAG}, Leu_{UAA}, Gln_{UUG}, Ile_{GAU1}, Ile_{GAU2}, Tyr_{GUA}, Ile_{CAU}, Glu_{UUC}, Pro_{UGG}, and Trp_{CCA}) (Abad et al. 2014). The observed patterns of nucleotide changes indicate some preferences that may reflect the constraints of the underlying machinery, but the types of nucleotide changes that are observed as a consequence of editing reactions span the entire range of possible nucleotide changes when editing events that have been experimentally validated in amoebozoans are considered together (Table 7.2). However, the pattern of events observed in any individual species does not necessarily follow the overall trend, which further supports the independent acquisition of specific nucleotide changes in each mitochondrial genome and the existence of a broadly active suite of enzymes in each organism that is able to restore the correct mt-tRNA sequences in each case. For example, in *D. discoideum* the majority (73.3%) of editing events entail a nucleotide edited to a pyrimidine versus only 26.7% of editing events that entail a nucleotide edited to a purine (Table 7.2). In contrast, the dominant types of changes observed when all species are considered together are those that result in changes to a purine nucleotide as a consequence of editing. Similarly, in *D. discoideum* it is most common for a nucleotide to be edited to a C (46.7% of editing events), followed by being edited to U or G (each comprises 26.7% of editing events). Finally, editing to an A is not observed, whereas editing to

Table 7.2 Summary of overall nucleotide changes observed in experimentally validated amoebozoan mitochondrial mt-tRNA 5′-editing events

	A-to-G	G-to-A	U-to-C	C-to-U	U-to-A	A-to-U	U-to-G	G-to-U	A-to-C	C-to-A	C-to-G	G-to-C	N-to-A	N-to-G	N-to-C	N-to-U
	Transition				Transversion											
Total	11	0	5	1	3	4	12	0	8	1	3	0	4	26	13	5
	(22.9)	*(0.0)*	*(10.4)*	*(2.1)*	*(6.3)*	*(8.3)*	***(25.0)***	*(0.0)*	*(16.7)*	*(2.1)*	*(6.3)*	*(0.0)*	*(8.3)*	***(54.2)***	*(27.1)*	*(10.4)*
D. discoideum	3	0	2	1	0	3	1	0	5	0	0	0	0	4	7	4
	(20.0)	*(0.0)*	*(13.3)*	*(6.7)*	*(0.0)*	*(20.0)*	*(6.7)*	*(0.0)*	***(33.3)***	*(0.0)*	*(0.0)*	*(0.0)*	*(0.0)*	*(26.7)*	***(46.7)***	*(26.7)*

	A editing	G editing	C editing	U editing	Pyrimidine to Purine	Purine to Pyrimidine	Purine to Purine	Pyrimidine to Pyrimidine	N to Purine	N to Pyrimidine
Total	23	0	5	20	19	12	11	6	30	18
	(47.9)	*(0.0)*	*(10.4)*	*(41.7)*	***(39.6)***	*(25.0)*	*(22.9)*	*(12.5)*	***(62.5)***	*(37.5)*
D. discoideum	11	0	1	3	1	8	3	3	4	11
	(73.3)	*(0.0)*	*(6.7)*	*(20.0)*	*(6.7)*	***(53.3)***	*(20.0)*	*(20.0)*	*(26.7)*	***(73.3)***

Italicized numbers in parentheses denote % of each type of editing event in the different categories listed (numbers in bold represent the most prominent % in each category). A editing, G editing, etc.: editing events where the indicated nucleotide (A, G, C, or U) is converted to a different nucleotide as a consequence of the editing reaction

A nucleotides comprises a substantial fraction (8.3%) of overall amoebozoan 5-′-editing events presented here (Table 7.2). One trend in which *D. discoideum* editing patterns match the average editing patterns is in terms of the nucleotide that is most commonly changed by 5′-editing. In this case, A is the most commonly edited nucleotide (overwhelmingly so in *D. discoideum*), while editing of other nucleotides is less common, in all species investigated to date (Table 7.2).

7.3 Role of TLPs in the Mechanism of 5′-Editing

The analysis of mt-tRNA 5′-editing in *D. discoideum* set the stage for characterization of the enzymatic machinery involved in this process, since many well-established genetic and biochemical tools have been developed for this organism. Sequencing results suggested that the overall mechanism for mt-tRNA 5′-editing in *D. discoideum* is analogous to that in *A. castellanii*, and therefore the relevant enzymes were expected to carry out a two-step process entailing posttranscriptional excision of mismatched nucleotides from the 5′-end followed by restoration of WC base pairing to the 5′-truncated tRNA by a polymerase acting in the 3′-to-5′ direction (Lonergan and Gray 1993a, b; Price and Gray 1999a, b; Bullerwell and Gray 2005; Jackman et al. 2012; Abad et al. 2014; Long and Jackman 2015) (Fig. 7.1). Consistent with this idea, four distinct gene-encoding members of the Thg1/TLP superfamily of enzymes were identified in the genome of *D. discoideum*. This finding was particularly intriguing in light of a recent biochemical characterization of a bacterial TLP that was able to exploit a template-dependent 3′-to-5′ polymerase activity to carry out 5′-end repair of a model tRNA substrate in vitro (Rao et al. 2011), suggesting that one or more of the *D. discoideum* TLPs could be acting similarly as the 3′-to-5′ polymerase component of the editing enzyme.

Biochemical and genetic tools were used to demonstrate that indeed one of the four encoded Thg1/TLP genes in *D. discoideum* is the enzyme responsible for the editing reaction in this organism. The four *D. discoideum* genes consist of one Thg1 enzyme (DdiThg1) and three TLPs (DdiTLP2–4), which are distinguished based on characteristic sequence differences (Jackman et al. 2012). In vitro characterization of the four purified recombinant enzymes with model mt-tRNA substrates demonstrated that two of the enzymes, DdiTLP3 and DdiTLP4, both exhibited robust in vitro activity with a variety of 5′-truncated mt-tRNAs that mimic editing intermediates in this pathway, intimating a possible in vivo role in 5′-editing for either or both of these proteins (Abad et al. 2011; Long and Jackman 2015; Long et al. 2016). However, of these two candidate editing enzymes, only DdiTLP3 was also localized to the mitochondria and exhibited a defect in the extent of mt-tRNA 5′-editing when its expression was depleted by RNAi, unambiguously demonstrating that it has an exclusive role in the 5′-editing reaction (Long et al. 2016). Notably, a viable *D. discoideum* strain with a deletion of DdiTLP3 could not be obtained, and the RNAi depletion strain shows significant growth and developmental defects. These defects are consistent with an essential requirement for 5′-editing of mt-tRNAs in

mitochondria of *D. discoideum* and the critical role for DdiTLP3 in this process. Subsequently, two genes encoding TLPs (AcaTLP1 and AcaTLP2) were identified in *A. castellanii*, with AcaTLP2 having a mitochondrial targeting sequence and likely to represent the 3′-to-5′ nucleotidyltransferase component of the mt-tRNA 5′-editing system in this organism (Rao et al. 2013).

Biochemical characterization of the 5′-editing enzyme DdiTLP3 suggests that the polymerase components of the editing enzyme are likely to be key players in the organism-specific rules regarding retention or removal of wobble base pairs, as described above. A kinetic comparison of the *D. discoideum* vs. *A. castellanii* 3′-to-5′ polymerases with some model mt-tRNA substrates revealed that the biochemical properties of these enzymes match the patterns of editing that are observed in their respective organisms (Long and Jackman 2015). Specifically, the 5′-editing TLP in *D. discoideum* (DdiTLP3) cannot efficiently extend (in the 3′-to-5′ direction) a mt-tRNA intermediate in which the 5′-terminal nucleotide is involved in a wobble base pair, suggesting that this type of base pair must be removed from the tRNA (presumably by the nuclease activity) prior to efficient repair of the mt-tRNA. Indeed, this is also consistent with the biological editing pattern observed in the organism. In contrast, neither of the putative editing TLPs in *A. castellanii* exhibits this strong distinction between repair of WC-terminated vs. wobble-terminated mt-tRNA intermediates, consistent with the ability of the 5′-editing pathway in this organism to retain the wobble base pair immediately adjacent to the site of 5′-editing in at least one mature mt-tRNA (Ala_{UGC}) in this species. More complex patterns of partial editing that have been observed in other species likely represent additional substrate recognition patterns that are yet to be revealed. A complete understanding of these will require a broader investigation of the biochemical properties of TLPs (and other parts of the editing machinery) from diverse species.

Interestingly, the identity of the nuclease involved in the first step of the 5′-editing pathway still remains unknown in any organism (Abad et al. 2011; Long and Jackman 2015; Long et al. 2016). Initial labeling experiments of mt-tRNAs in purified *A. castellanii* mitochondrial extracts led to the hypothesis that a 5′-exonuclease or endonuclease is responsible for the first step of mt-tRNA 5′-editing (Price and Gray 1999b). Potential 5′-to-3′ exonuclease candidates include enzymes such as Rat1 or Xrn1, which are known to act on tRNA during the rapid tRNA decay (RTD) pathway that removes incorrectly modified or processed tRNA from eukaryotic cells (Whipple et al. 2011; Dewe et al. 2012; Betat et al. 2014). However their high degree of processivity would seem to preclude these particular nucleases from removing only one to three nucleotides from the 5′-end to generate the relevant editing intermediates that have been observed experimentally. The 5′-end-associated endonuclease activity of RNase P, normally used to remove additional 5′ leader sequences from primary tRNA transcripts, is also a candidate (Yuan and Altman 1995). However the characteristic precision in cleavage site selection exhibited by RNase P also makes it unlikely that this enzyme is involved in a process such as 5′-editing that requires cleavage at a variety of positions within the aminoacyl acceptor stem (Yuan and Altman 1995). An unknown 5-′-endonuclease may recognize and cleave between mismatches. However in light

of the differing tolerances of G∘U editing between different mt-tRNAs in *D. discoideum* as well as *A. castelllani*, a specialized mechanism of recognition would be required to specifically target some G∘U base pairs but not others. Finally it is possible that the TLPs that perform the repair reaction also possess a nucleolytic ability; however, further study is required to address this possibility.

7.4 Other Mitochondrial Functions for 3′-to-5′ Polymerases

Analysis of the remaining three Thg1/TLP enzymes encoded by *D. discoideum* demonstrated that each of the four enzymes specified by this organism appears to exhibit distinct and non-overlapping physiological roles. For two of the enzymes (DdiThg1 and DdiTLP4), these appear to be strictly cytoplasmic roles involving nucleus-encoded RNAs, although the actual biological function is only known for DdiThg1, which adds G_{-1} to $tRNA^{His}$ like other eukaryotic Thg1 orthologs. The ability of DdiTLP4 to act on at least two different noncoding RNA substrates raises the possibility that this enzyme is involved in the broader metabolism of noncoding RNAs, but a definitive role for this enzyme in RNA metabolism remains to be determined (Long et al. 2016). However, DdiTLP2, like DdiTLP3, localizes to mitochondria and is the only one of the four DdiTLPs for which a viable deletion strain could be obtained; thus, its activity is not strictly essential for growth in *D. discoideum*. Sequence analysis of the mitochondrion-encoded $tRNA^{His}$ obtained from DdiTLP2 deletion strains revealed a total absence of the G_{-1} nucleotide compared to the same tRNA isolated from the parental strain, where the G_{-1} nucleotide could readily be observed (Long et al. 2016). Biochemical analysis of the purified DdiTLP2 enzyme was consistent with this role, since DdiTLP2 catalyzed robust incorporation of the G_{-1} nucleotide into mt-$tRNA^{His}$ transcripts. This in vivo and in vitro evidence confirmed that DdiTLP2 plays a role in $tRNA^{His}$ maturation that is distinct from the traditional 5′-editing reaction but is related in the sense that an additional unencoded nucleotide is added to the 5′-end of this tRNA. Intriguingly, DdiTLP3 was also observed to catalyze this G_{-1} addition activity in vitro, but the absence of any detectable G_{-1} on the isolated mt-$tRNA^{His}$ in the *dditlp2* deletion strains (in which DdiTLP3 is intact) suggests that this overlap in function is not similarly observed in the context of the mitochondria themselves.

The observation of a $tRNA^{His}$-related function for DdiTLP2 is highly reminiscent of the activity exhibited by the Thg1 members of the enzyme family, whose essential function in nucleus-encoded $tRNA^{His}$ metabolism through addition of the G_{-1} nucleotide is well-established (Gu et al. 2005; Preston and Phizicky 2010). However, there are some significant biochemical differences between the DdiThg1-catalyzed and DdiTLP2-catalyzed activities that suggest that these may have arisen independently. First, the eukaryotic G_{-1} addition function generates a non-WC base pair (G_{-1} is added opposite A_{73}), while in *D. discoideum* mitochondria, the presence of a C_{73} discriminator nucleotide would allow the TLP to take advantage of its preference for generating WC base pairs, as observed for all TLPs to date (Abad et al. 2010; Rao

et al. 2011). Second, selective recognition of $tRNA^{His}$ as a substrate for Thg1 is strictly dependent on the presence of the $tRNA^{His}$ GUG anticodon, while DdiTLP2 readily acts on mt-$tRNA^{His}$ variants in which the GUG is altered to another anticodon sequence (Jackman and Phizicky 2006; Long et al. 2016). Interestingly, many metazoan species encode a Thg1 enzyme with a putative N-terminal mitochondrial targeting peptide. In fact, the human THG1-like (THG1L) enzyme has been localized to mitochondria, where its overexpression is associated with increased mitochondrial proliferation that may be relevant to the progression of diabetic nephropathy (Murphy et al. 2008, 2013; Hickey et al. 2011; Corcoran et al. 2013). However, the precise nature of the biochemical activities of THG1L in human mitochondria is not known. A complete understanding of all mitochondrial roles for members of the Thg1/TLP enzyme family will require a much better understanding of the molecular functions of these unusual enzymes.

7.5 Origin and Evolution of Mitochondrial 5′-Editing

How do RNA editing systems originate, and why do they persist? The initial examples of this phenomenon raised these questions, and they have preoccupied researchers in this field for more than three decades as novel RNA editing systems have been uncovered. A common perception is that editing evolved to correct "mistakes" in RNA transcripts, resulting from what would otherwise be deleterious or even lethal mutations in the corresponding gene. Indeed, the function of mt-tRNA 5′-editing is effectively to restore the helical character of a mismatched acceptor stem, the integrity of which plays an important role in determining aminoacylation specificity (Giegé et al. 1998), as well as being essential in the formation of the tertiary structure of tRNA, thereby positioning the attached amino acid correctly for peptide bond synthesis (Kim 1978).

However, an explanation for the emergence of an RNA editing system based on its current function puts the cart before the horse, so to speak, because deleterious or lethal mutations are not likely to persist in a genome while waiting for a corrective RNA editing system to evolve. Such mutations would be eliminated by purifying selection long before the appearance of the required editing system. Although, undoubtedly, the various RNA editing systems that have been described *currently* serve a corrective function, we would argue that they did not emerge initially *in response to* a need for editing.

To explain the origin of RNA editing, a three-stage, neutral evolutionary model (Covello and Gray 1993), subsequently termed "constructive neutral evolution (CNE)" (Stoltzfus 1999; Gray et al. 2010; Lukeš et al. 2011), has been proposed. The first stage of this model posits that, whatever the nature of the editing activity, it was already in place *before* there was a need for editing, emerging in a neutral manner (i.e., not subject to selection) from a preexisting activity in the cell. The second stage envisages mutation at editable sites, mutations that are able to be acted upon at the RNA level and effectively "corrected" by the preexisting editing system.

In the case of tRNA 5′-editing, editable sites are the first three positions of the tRNA. Finally, in the third stage, as the number of sites requiring editing increases and the probability of complete reversion to a non-editing-requiring state becomes vanishing small [a Muller's ratchet type of effect (Muller 1964)], selection comes into play, with the RNA editing system becoming an essential part of the genetic information transfer pathway. An important corollary is that the appearance of an RNA editing system is in itself mutagenic, because it allows the fixation in the genome of otherwise deleterious/lethal mutations that would normally be eliminated by purifying selection (Gray 2001).

In the specific case of mt-tRNA 5′-editing, the initial discovery of this process in *A. castellanii* (Lonergan and Gray 1993a, b) immediately posed a conundrum. Assuming that the ancestral state was one in which the tRNA acceptor stem was fully paired, how did most of the mtDNA-encoded tRNAs (12/15) in *A. castellanii* come to require editing (i.e., those tRNAs that display one or more mismatches within the first three base pairs of the acceptor stem)? And, why are mismatches strictly confined to the first three base pairs of the acceptor stem, the rest of the tRNA assuming a conventional structure? More precisely, since we know that editing occurs on the 5′ side of the acceptor stem (Lonergan and Gray 1993a), why are only the first three tRNA positions so evolutionarily labile?

Invoking a CNE model as an explanation, we envisage as the first stage the appearance in mitochondria of a constitutive system that is able to remove the first three nucleotides at the 5′-end of a tRNA and replace those nucleotides using the corresponding positions on the 3′ side of the acceptor stem as a template, generating standard WC-type base pairs as a result. (How such a system comes to be lodged in mitochondria in the first instance is addressed below.) The postulated RNA editing system comprises both a nuclease, either exo- or endo-, capable of carrying out the first step—removal of 5′ nucleotides—and a novel, template-directed 3′-to-5′ nucleotidyltransferase acting sequentially: essentially a 3′-to-5′ RNA-dependent RNA polymerase. As discussed earlier, these two activities may reside in separate proteins, or in a single protein. A mitochondrial system having these properties was, in fact, biochemically characterized in *A. castellanii* (Price and Gray 1999b), as well as in *Spizellomyces punctatus* (Bullerwell and Gray 2005), a chytrid fungus in which tRNA 5′-editing had also been described (Laforest et al. 1997).

Once such a system appears in mitochondria, it is straightforward to envisage the consequences. Mutations that disrupt base pairing in the first three positions of the acceptor stem would preferentially accumulate because these mutations can be counteracted at the level of the tRNA transcripts. In essence, relaxed functional constraints at the genome level would allow the fixation of mutations that would otherwise be purged by purifying selection. Assuming that the editing activity is able to operate on any mitochondrial tRNA, a second consequence is the spread of tRNA 5′ mismatching among the collection of mtDNA-encoded tRNAs; most tRNAs would be expected to exhibit at least one 5′ mismatch, but a few (by chance) might have none. Finally, because the mutational process resulting in tRNA 5′ mismatching is stochastic, we expect that there should be no conservation of editing

sites in related organisms in which this type of editing occurs: i.e., we attribute no functional significance to the *emergence* of mismatching.

Consideration of characterized editing sites in *A. castellanii* mitochondrial tRNAs and sequence comparisons among mitochondrial tRNA isoacceptors within different *Acanthamoeba* species provides strong evidence for the second stage of the CNE model: mutation at editable sites. In *A. castellanii*, a major component of mitochondrial genome evolution is AT drift whereby both coding but particularly noncoding regions are enriched in A and T nucleotides (70.6% overall; 71.6% in non-coding regions). Of 23 positions known to be edited in 12 mitochondrial tRNAs, 20 (87%) represent a change from A or U to another nucleotide (mostly G), indicating that the first three 5′ positions, left to accumulate mutations without functional constraint, become even substantially more A + T-rich than the mitochondrial genome as a whole. A similar pattern is evident in the case of *D. discoideum* (Table 7.2).

When mitochondrial tRNA sequences for the same isoacceptor are compared from different *Acanthamoeba* species (Fig. 7.2), it is evident that the first three 5′ positions (i.e., the editable positions) are the only ones that differ substantially; moreover, these variant positions are predominantly AT-rich, as expected. With few exceptions, the sequence at all other tRNA positions is identical, in particular at the final four 5′ positions and all seven of the 3′ positions of the acceptor stem.

Comparison of acceptor stems for the same mitochondrial isoacceptor tRNA in different amoebozoan genera demonstrates that the pattern of editing is not conserved at any of the first three 5′ positions (Fig. 7.3). For example, in $tRNA^{Ile}_{CAU}$, as predicted from the corresponding gene sequence, the first acceptor stem pairing is G∘U in *Veramoeba vermiformis* (*Vve*; a potential G→A editing), A×C in *D. discoideum* (*Ddi*; a demonstrated A→G editing), A×G in *D. citrinum* (*Dci*; a potential A→C editing), but A-U in *A. castellanii* (*Aca*) and G-C in *D. fasciculatum* (*Dfa*) and in three isolates of *P. pallidum* (*Ppa*; no editing required). Notably, non-conservation of editing sites is clearly evident among different species of the same genus (*Dictyostelium*). These observations are consistent with the view that the emergence of editing sites is stochastic and not related to function (i.e., the emergence is neutral, not selected).

How can we account for the appearance of the editing system in mitochondria, even before there is a need for editing? The occurrence of mt-tRNA 5′-editing shows a punctuate distribution: i.e., eukaryotes carrying out this type of editing are interspersed with related species in which there is no obvious requirement for mt-tRNA 5′-editing (Jackman et al. 2012). This pattern makes vertical inheritance of the system from a remote common ancestor difficult to envisage, if in fact the majority of eukaryotes seem not to engage in mt-tRNA 5′-editing. The examples identified so far of this type of editing more likely represent derived traits within the specific lineages in which they occur. If so, how then do we explain the virtually identical biochemical characteristics of partially purified mt-tRNA 5′-editing systems from two distantly related eukaryotes, the amoebozoan *A. castellanii* (Price and Gray 1999b) and the chytrid fungus *S. punctatus* (Bullerwell and Gray 2005): microbes that represent two distinct eukaryotic supergroups?

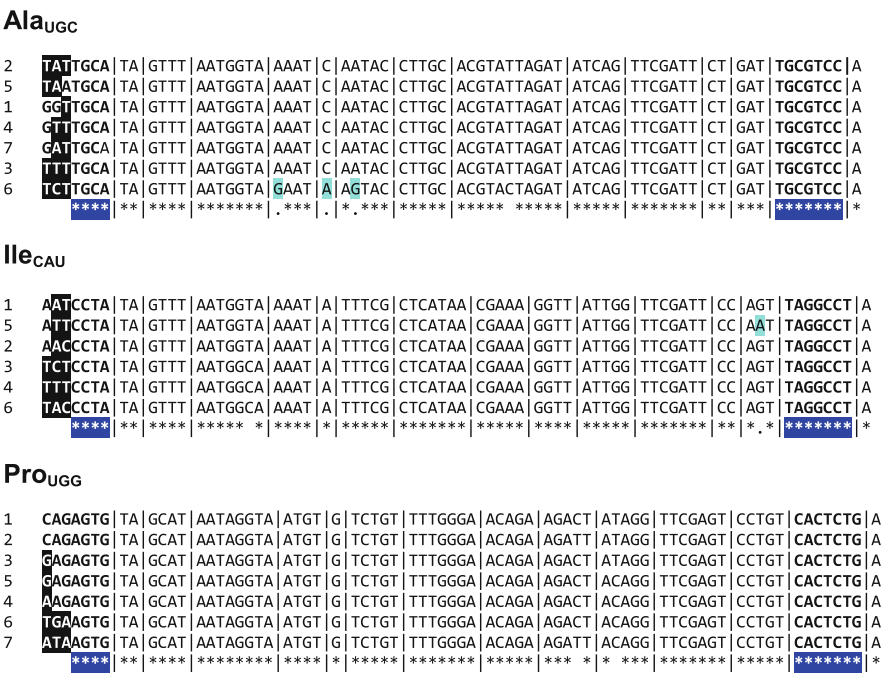

Fig. 7.2 *Acanthamoeba* sp. mitochondrial tRNA isoacceptors are essentially identical in sequence *except* at the first three 5′ positions of the acceptor stem. Sequence data (tDNA) are from Ledee and Byers (2009) and are aligned for three different tRNAs. Mismatched positions (nucleotides on the 5′ side of the acceptor stem that do not pair with the corresponding nucleotide on the 3′ side) are denoted by white letters on a black background. Nucleotides at each position that are identical among the compared isoacceptors are denoted by asterisks (*) below the alignment; those identical nucleotides in the acceptor stem are shown as white asterisks on a blue background. The few non-identical nucleotides (denoted by a period below the alignment) are highlighted in turquoise. Numbers indicate different *Acanthamoeba* species

As discussed earlier, studies in *D. discoideum* (Abad et al. 2011, 2014) and *A. castellanii* (Rao et al. 2011) have provided compelling evidence that mitochondrion-targeted TLPs constitute the 3′-to-5′ nucleotidyltransferase that carries out the resynthesis step in the mt-tRNA 5′-editing pathway. TLPs are broadly distributed in all three domains of life—Archaea, Bacteria, and Eukarya—with bacterial TLPs constituting two distinct clades, group 1 and group 2, in phylogenetic trees (Heinemann et al. 2010; Jackman et al. 2012). Eukaryotic TLPs appear monophyletic and are specifically associated with group 1 bacterial TLPs, suggesting an origin from the latter group as a result of one or more horizontal gene transfers during eukaryotic evolution. A striking correlation exists between the presence of a putatively mitochondrion-targeted TLP and the existence of mt-tRNA 5′-editing, either demonstrated or inferred from the secondary structure of mtDNA-encoded tRNAs (Table 7.3).

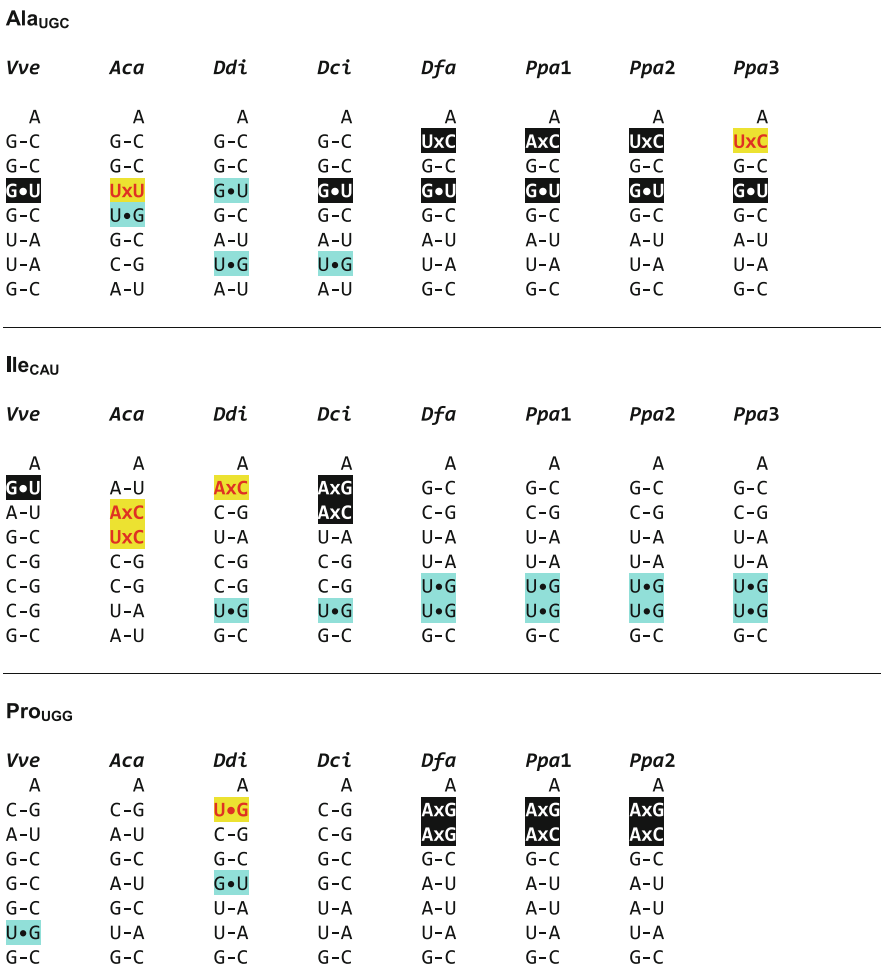

Fig. 7.3 Lack of conservation of editing sites in mitochondrial tRNA isoacceptors from different amoebozoan genera. Confirmed editing sites are shown in red on a yellow background. Predicted editing sites are indicated by white lettering on a black background. G∘U and U∘G pairs that are NOT subject to editing within the acceptor stem are highlighted in turquoise. Abbreviations: *Vve Veramoeba vermiformis*, *Aca Acanthamoeba castellanii*, *Ddi Dictyostelium discoideum*, *Dci Dictyostelium citrinum*, *Dfa Dictyostelium fasciculatum*, *Ppa*1 *Polysphondylium pallidum* CK8, *Ppa*2 *Polysphondylium pallidum* PN500, *Ppa*3 *Polysphondylium pallidum* PPHU8

Given these observations, we propose the following scenario for the *independent* acquisition of biochemically highly similar mt-tRNA 5′-editing systems in select eukaryotic lineages. We posit that a group 1 bacterial-type TLP was acquired by horizontal gene transfer early in eukaryotic cell evolution. The original function of such a protein is not clear, although we have suggested that, in general, TLPs may act constitutively to reconstitute (in a 3′-to-5′ direction) 5′ ends that suffer loss of nucleotides, e.g., due to stochastic 5′-3′ exonucleolysis: in much the same way as

Table 7.3 Predicted and verified mt-tRNA 5′-editing among eukaryotes

	mtDNA[a]	tRNAs[b]	Predicted[c]	Confirmed[d]	mt-TLP[e]	Reference
Amoebozoa						
Acanthamoeba castellanii	U12386	15	12	12	AFS33774	Lonergan and Gray (1993a, b), Price and Gray (1999a, b)
Dictyostelium discoideum	NC_000895	18	10	10	XP_641524	Abad et al. (2011, 2014)
Dictyostelium fasciculatum	EU275727	17	10 (+1*)		XP_004361357	
Dictyostelium citrinum	DQ336395	19	6 (+6*)		n.d.	
Polysphondylium pallidum CK8	AY700145	18	13		n.d.	
Polysphondylium pallidum PN500	EU275726	20	10 (+1*)	5	XP_020430217[f]	Abad et al. (2014)
Physarum polycephalum	NC_002508	5	2	2	[g]	Gott et al. (2010)
Veramoeba vermiformis	GU828005	25	12*		n.d.	
Phalansterium sp.	KC121006	24	15 (+2*)		n.d.	Pombert et al. (2013)
Paramoeba pemaquidensis	KX611830	19	15 (+2*)		n.d.	Tanifuji et al. (2017)
Fungi						
Spizellomyces punctatus	NC_003052	8	8	6	SPPG_03437	Laforest et al. (1997)
Harpochytrium sp. JEL94	AY182005	8	6	2	n.d.	Bullerwell et al. (2003)
Harpochytrium sp. JEL105	AY182006	8	7	3	n.d.	
Monoblepharella sp. JEL15	AY182007	9	7	4	n.d.	
Hyaloraphidium curvatum	NC_003048	7	3		n.d.	Forget et al. (2002)
Cercozoa						
Bigelowiella natans	HQ840955	26	4 (+7*)[h]		[i]	Tanifuji et al. (2016)
Lotharella oceanica	KT806043	24	1 (+7*)		n.d.	

(continued)

Table 7.3 (continued)

	mtDNA[a]	tRNAs[b]	Predicted[c]	Confirmed[d]	mt-TLP[e]	Reference
Heterolobosea						
Naegleria gruberi	NC_002573	21	8 (+1*)		XP_002683372	
Naegleria fowleri	NC_021104	21	7 (+2*)		n.d.	Herman et al. (2013)
Acrasis kona	NC_026286	13	6 (+2*)		n.d.	Fu et al. (2014)

[a]GenBank accession numbers for complete mitochondrial genome sequences from which mt-tRNA 5′-editing has been or can be inferred
[b]Total number of tRNAs encoded by the mitochondrial genome in question
[c]Number of tRNAs predicted to undergo 5′-editing (numbers with asterisks indicate tRNAs having only U∘G/ G∘U wobble pairs within the first three positions of the acceptor stem)
[d]Number of tRNAs in which 5′-editing has been experimentally confirmed
[e]GenBank accession number (*n.d.* not determined)
[f]Single TLP identified in *P. pallidum* PN500 is predicted to lack an N-terminal mitochondrial targeting sequence. Corresponding genome sequence encodes a 65-aa N-terminal extension containing a predicted mitochondrial targeting signal
[g]A predicted full-length mitochondrion-targeted TLP sequence can be deduced from genomic and transcriptomic data available at http://www.physarum-blast.ovgu.de/ (unpublished results)
[h]Four additional tRNAs have mismatches only further down the acceptor stem (three at position 4, one at position 7)
[i]*B. natans* genome project. Joint Genome Institute (complete TLP sequence available at http://genome.jgi-psf.org/cgi-bin/dispGeneModel?db=Bigna1&id=82370)

the well-known CCA-adding enzyme (tRNA nucleotidyltransferase) serves to reconstitute the 3′ ends of tRNAs, but in a 5′-to-3′ direction. TLP gene duplication in selected eukaryotes followed by acquisition by one of the duplicates of an N-terminal mitochondrial import sequence would target a TLP to mitochondria, setting in motion the CNE model discussed earlier, in which mtDNA-encoded tRNA genes gradually accumulate mispairing mutations at positions 1, 2, and/or 3 that are able to be "corrected" at the RNA level. This scenario accounts both for the punctate distribution of mt-tRNA 5′-editing (due to independent emergence of a mitochondrion-targeted TLP in different lineages) and the high degree of biochemical similarity in distantly related eukaryotes (because although these systems are independently derived traits, they are underpinned by orthologous TLPs).

Note that implicit in this scenario is the assumption that TLPs are only able to "repair" the first three nucleotides at the 5′-end of a tRNA, in much the same way that the CCA-adding enzyme adds a maximum of three nucleotides at the 3′ end. This assumption is necessary to account for the fact that mismatch-inducing mutations are restricted to the first three positions of the acceptor stem: non-repairable mutations further down the acceptor stem presumably being eliminated by purifying selection. This assumption remains to be tested biochemically, however.

7.6 Conclusion

The discovery of RNA editing systems provides an inherent challenge to the principle of the genome serving as the sole information carrier for the cell. In this example of mitochondrial 5′-editing, clearly the editing enzymes are essential partners in helping to assemble the complete set of tRNA species that are required for translation. Given the predicted bacterial origin of the TLP components of the editing machinery, it is intriguing to consider the possibilities for similar roles of these enzymes in extant bacterial and archaeal species where they are encoded in the genome but where 5′-editing is not predicted to occur. Importantly, since only a limited number of TLP enzymes have been investigated biochemically and genetically, and since the relevant editing nuclease has yet to be identified in any species, the potential for these enzymes to participate more broadly in the maintenance or repair of cellular nucleic acids will be exciting to evaluate through future studies of this intriguing enzyme family.

References

Abad MG, Rao BS, Jackman JE (2010) Template-dependent 3′–5′ nucleotide addition is a shared feature of $tRNA^{His}$ guanylyltransferase enzymes from multiple domains of life. Proc Natl Acad Sci U S A 107:674–679

Abad MG, Long Y, Willcox A, Gott JM, Gray MW, Jackman JE (2011) A role for $tRNA^{His}$ guanylyltransferase (Thg1)-like proteins from *Dictyostelium discoideum* in mitochondrial 5′-tRNA editing. RNA 17:613–623

Abad MG, Long Y, Kinchen RD, Schindel ET, Gray MW, Jackman JE (2014) Mitochondrial tRNA 5′-editing in *Dictyostelium discoideum* and *Polysphondylium pallidum*. J Biol Chem 289:15155–15165

Benne R, Van den Burg J, Brakenhoff JP, Sloof P, Van Boom JH, Tromp MC (1986) Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. Cell 46:819–826

Betat H, Long Y, Jackman JE, Mörl M (2014) From end to end: tRNA editing at 5′- and 3′-terminal positions. Int J Mol Sci 15:23975–23998

Bullerwell CE, Gray MW (2005) *In vitro* characterization of a tRNA editing activity in the mitochondria of *Spizellomyces punctatus*, a chytridiomycete fungus. J Biol Chem 280:2463–2470

Bullerwell CE, Forget L, Lang BF (2003) Evolution of monoblepharidalean fungi based on complete mitochondrial genome sequences. Nucleic Acids Res 31:1614–1623

Burger G, Plante I, Lonergan KM, Gray MW (1995) The mitochondrial DNA of the amoeboid protozoon, *Acanthamoeba castellanii*: complete sequence, gene content and genome organization. J Mol Biol 245:522–537

Chen S-H, Habib G, Yang C-Y, Gu Z-W, Lee B, Weng S-a, Silberman CS-J, Deslypere J, Rosseneu M, Gotto A, Li W-H, Chan L (1987) Apolipoprotein B-48 is the product of a messenger RNA with an organ-specific in-frame stop codon. Science 238:363–366

Cooley L, Appel B, Söll D (1982) Post-transcriptional nucleotide addition is responsible for the formation of the 5′ terminus of histidine tRNA. Proc Natl Acad Sci U S A 79:6475–6479

Corcoran JB, McCarthy S, Griffin B, Gaffney A, Bhreathnach U, Börgeson E, Hickey FB, Docherty NG, Higgins DF, Furlong F, Martin F, Godson C, Murphy M (2013) IHG-1 must be localised to mitochondria to decrease Smad7 expression and amplify TGF-β1-induced fibrotic responses. Biochim Biophys Acta 1833:1969–1978

Covello PS, Gray MW (1993) On the evolution of RNA editing. Trends Genet 9(8):265–268

Dewe JM, Whipple JM, Chernyakov I, Jaramillo LN, Phizicky EM (2012) The yeast rapid tRNA decay pathway competes with elongation factor 1A for substrate tRNAs and acts on tRNAs lacking one or more of several modifications. RNA 18:1886–1896

Eichinger L, Pachebat JA, Glockner G, Rajandream MA, Sucgang R, Berriman M, Song J, Olsen R, Szafranski K, Xu Q, Tunggal B, Kummerfeld S, Madera M, Konfortov BA, Rivero F, Bankier AT, Lehmann R, Hamlin N, Davies R, Gaudet P, Fey P, Pilcher K, Chen G, Saunders D, Sodergren E, Davis P, Kerhornou A, Nie X, Hall N, Anjard C, Hemphill L, Bason N, Farbrother P, Desany B, Just E, Morio T, Rost R, Churcher C, Cooper J, Haydock S, van Driessche N, Cronin A, Goodhead I, Muzny D, Mourier T, Pain A, Lu M, Harper D, Lindsay R, Hauser H, James K, Quiles M, Madan Babu M, Saito T, Buchrieser C, Wardroper A, Felder M, Thangavelu M, Johnson D, Knights A, Loulseged H, Mungall K, Oliver K, Price C, Quail MA, Urushihara H, Hernandez J, Rabbinowitsch E, Steffen D, Sanders M, Ma J, Kohara Y, Sharp S, Simmonds M, Spiegler S, Tivey A, Sugano S, White B, Walker D, Woodward J, Winckler T, Tanaka Y, Shaulsky G, Schleicher M, Weinstock G, Rosenthal A, Cox EC, Chisholm RL, Gibbs R, Loomis WF, Platzer M, Kay RR, Williams J, Dear PH, Noegel AA, Barrell B, Kuspa A (2005) The genome of the social amoeba *Dictyostelium discoideum*. Nature 435:43–57

Forget L, Ustinova J, Wang Z, Huss VAR, Lang BF (2002) *Hyaloraphidium curvatum:* a linear mitochondrial genome, tRNA editing, and an evolutionary link to lower fungi. Mol Biol Evol 19:310–319

Fu C-J, Sheikh S, Miao W, Andersson SGE, Baldauf SL (2014) Missing genes, multiple ORFs, and C-to-U type RNA editing in *Acrasis kona* (Heterolobosea, Excavata) mitochondrial DNA. Genome Biol Evol 6:2240–2257

Giegé R, Sissler M, Florentz C (1998) Universal rules and idiosyncratic features in tRNA identity. Nucleic Acids Res 26:5017–5035

Gott JM, Somerlot BH, Gray MW (2010) Two forms of RNA editing are required for tRNA maturation in *Physarum* mitochondria. RNA 16:482–488

Gray MW (2001) Speculations on the origin and evolution of editing. In: Bass BL (ed) RNA editing, vol 34. Frontiers in molecular biology. Oxford University Press, Oxford, pp 160–184

Gray MW (2003) Diversity and evolution of mitochondrial RNA editing systems. IUBMB Life 55:227–233

Gray MW (2012) Evolutionary origin of RNA editing. Biochemistry 51:5235–5242

Gray MW, Lukeš J, Archibald JM, Keeling PJ, Doolittle WF (2010) Irremediable complexity? Science 330:920–921

Gu W, Jackman JE, Lohan AJ, Gray MW, Phizicky EM (2003) $tRNA^{His}$ maturation: an essential yeast protein catalyzes addition of a guanine nucleotide to the 5′ end of $tRNA^{His}$. Genes Dev 17:2889–2901

Gu W, Hurto RL, Hopper AK, Grayhack EJ, Phizicky EM (2005) Depletion of *Saccharomyces cerevisiae* $tRNA^{His}$ guanylyltransferase Thg1p leads to uncharged $tRNA^{His}$ with additional m^5 C. Mol Cell Biol 25:8191–8201

Heinemann IU, Randau L, Tomko RJJ, Söll D (2010) 3′-5′ $tRNA^{His}$ guanylyltransferase in bacteria. FEBS Lett 584:3567–3572

Herman EK, Greninger AL, Visvesvara GS, Marciano-Cabral F, Dacks JB, Chiu CY (2013) The mitochondrial genome and a 60-kb nuclear DNA segment from *Naegleria fowleri*, the causative agent of primary amoebic meningoencephalitis. J Eukaryot Microbiol 60:179–191

Hickey FB, Corcoran JB, Docherty NG, Griffin B, Bhreathnach U, Furlong F, Martin F, Godson C, Murphy M (2011) IHG-1 promotes mitochondrial biogenesis by stabilizing PGC-1α. J Am Soc Nephrol 22:1475–1485

Jackman JE, Phizicky EM (2006) $tRNA^{His}$ guanylyltransferase adds G_{-1} to the 5′ end of $tRNA^{His}$ by recognition of the anticodon, one of several features unexpectedly shared with tRNA synthetases. RNA 12:1007–1014

Jackman JE, Gott JM, Gray MW (2012) Doing it in reverse: 3′-to-5′ polymerization by the Thg1 superfamily. RNA 18:886–899

Jahn D, Pande S (1991) Histidine tRNA guanylyltransferase from *Saccharomyces cerevisiae.* II. Catalytic mechanism. J Biol Chem 266:22832–22836

Kim S-H (1978) Three-dimensional structure of transfer RNA and its functional implications. Adv Enzymol Relat Areas Mol Biol 46:279–315

Laforest MJ, Roewer I, Lang BF (1997) Mitochondrial tRNAs in the lower fungus *Spizellomyces punctatus*: tRNA editing and UAG 'stop' codons recognized as leucine. Nucleic Acids Res 25:626–632

Laforest M-J, Bullerwell CE, Forget L, Lang BF (2004) Origin, evolution, and mechanism of 5′ tRNA editing in chytridiomycete fungi. RNA 10:1191–1199

Ledee DR, Byers TJ (2009) Length and sequence heterogeneity in the mitochondrial internal transcribed spacer of *Acanthamoeba* spp. J Eukaryot Microbiol 56:257–262

Lohan A, Gray M (2007) Analysis of 5′- or 3′-terminal tRNA editing: mitochondrial 5′ tRNA editing in Acanthamoeba castellanii as the exemplar. Methods Enzymol 424:223–242

Lonergan KM, Gray MW (1993a) Editing of transfer RNAs in *Acanthamoeba castellanii* mitochondria. Science 259:812–816

Lonergan KM, Gray MW (1993b) Predicted editing of additional transfer RNAs in *Acanthamoeba castellanii* mitochondria. Nucleic Acids Res 21:4402

Long Y, Jackman JE (2015) In vitro substrate specificities of 3′–5′ polymerases correlate with biological outcomes of tRNA 5′-editing reactions. FEBS Lett 589:2124–2130

Long Y, Abad MG, Olson ED, Carrillo EY, Jackman JE (2016) Identification of distinct biological functions for four 3′-5′ RNA polymerases. Nucleic Acids Res 44:8395–8406

Lukeš J, Archibald JM, Keeling PJ, Doolittle WF, Gray MW (2011) How a neutral evolutionary ratchet can build cellular complexity. IUBMB Life 63:528–537

Muller HJ (1964) The relation of recombination to mutational advance. Mutat Res 1:2–9

Murphy M, Docherty NG, Griffin B, Howlin J, McArdle E, McMahon R, Schmid H, Kretzler M, Droguett A, Mezzano S, Brady HR, Furlong F, Godson C, Martin F (2008) IHG-1 amplifies TGF-β1 signaling and is increased in renal fibrosis. J Am Soc Nephrol 19:1672–1680

Murphy M, Hickey F, Godson C (2013) IHG-1 amplifies TGF-β1 signalling and mitochondrial biogenesis and is increased in diabetic kidney disease. Curr Opin Nephrol Hypertens 22:77–84

Ogawa S, Yoshino R, Angata K, Iwamoto M, Pi M, Kuroe K, Matsuo K, Morio T, Urushihara H, Yanagisawa K, Tanaka Y (2000) The mitochondrial DNA of *Dictyostelium discoideum*: complete sequence, gene content and genome organization. Mol Gen Genet 263:514–519

Pombert J-F, Smirnov A, James ER, Janouškovec J, Gray MW, Keeling PJ (2013) The complete mitochondrial genome from an unidentified *Phalansterium* species. Protist Genom 1:25–32

Powell LM, Wallis SC, Pease RJ, Edwards YH, Knott TJ, Scott J (1987) A novel form of tissue-specific RNA processing produces apolipoprotein-B48 in intestine. Cell 50:831–840

Preston MA, Phizicky EM (2010) The requirement for the highly conserved G_{-1} residue of *Saccharomyces cerevisiae* $tRNA^{His}$ can be circumvented by overexpression of $tRNA^{His}$ and its synthetase. RNA 16:1068–1077

Price DH, Gray MW (1998) Editing of tRNA. In: Grosjean H, Benne R (eds) Modification and editing of RNA. ASM Press, Washington, pp 289–305

Price DH, Gray MW (1999a) Confirmation of predicted edits and demonstration of unpredicted edits in *Acanthamoeba castellanii* mitochondrial tRNAs. Curr Genet 35:23–29

Price DH, Gray MW (1999b) A novel nucleotide incorporation activity implicated in the editing of mitochondrial transfer RNAs in *Acanthamoeba castellanii*. RNA 5:302–317

Rao BS, Maris EL, Jackman JE (2011) tRNA 5′-end repair activities of $tRNA^{His}$ guanylyltransferase (Thg1)-like proteins from Bacteria and Archaea. Nucleic Acids Res 39:1833–1842

Rao BS, Mohammad F, Gray MW, Jackman JE (2013) Absence of a universal element for $tRNA^{His}$ identity in *Acanthamoeba castellanii*. Nucleic Acids Res 41:1885–1894

Stoltzfus A (1999) On the possibility of constructive neutral evolution. J Mol Evol 49:169–181

Tanifuji G, Archibald JM, Hashimoto T (2016) Comparative genomics of mitochondria in chlorarachniophyte algae: endosymbiotic gene transfer and organellar genome dynamics. Sci Rep 6:21016

Tanifuji G, Cenci U, Moog D, Dean S, Nakayama T, David V, Fiala I, Curtis BA, Sibbald SJ, Onodera NT, Colp M, Flegontov P, Johnson-MacKinnon J, McPhee M, Inagaki Y, Hashimoto T, Kelly S, Gull K, Lukeš J, Archibald JM (2017) Genome sequencing reveals metabolic and cellular interdependence in an amoeba-kinetoplastid symbiosis. Sci Rep 7:11688

Whipple J, Lane E, Chernyakov I, D'Silva S, Phizicky E (2011) The yeast rapid tRNA decay pathway primarily monitors the structural integrity of the acceptor and T-stems of mature tRNA. Genes Dev 25:1173–1184

Williams JB, Cooley L, Soll D (1990) Enzymatic addition of guanylate to histidine transfer RNA. Methods Enzymol 181:451–462

Yuan Y, Altman S (1995) Substrate recognition by human RNase P: identification of small, model substrates for the enzyme. EMBO J 14:159–168

Chapter 8
Editing of Mitochondrial RNAs in *Physarum polycephalum*

Jillian Houtz, Nicole Cremona, and Jonatha M. Gott

Abstract The mitochondrial transcriptome of the true acellular slime mold *Physarum polycephalum* (*Physarum*) undergoes extensive RNA editing to produce precise, site-specific changes not encoded at the DNA level. RNA editing in *Physarum* is essential for proper mitochondrial gene expression by creating open reading frames in protein-coding RNAs and by altering the folding stability of structural RNAs. *Physarum* carries out one of the widest range of RNA editing events yet described. These changes to mitochondrial RNAs involve the site-specific insertion of over 1300 "extra" nucleotides, deletion of 3 encoded nucleotides, targeted base conversions, and the removal and replacement of nucleotides at the 5′ end of certain tRNAs. While these sequence alterations are absolutely required for the production of functional transcripts, it remains a mystery why they are not encoded in the mitochondrial genome. Although various examples of RNA editing have been described in several eukaryotic organisms, *Physarum* mitochondria achieve non-templated nucleotide insertion by a unique co-transcriptional mechanism. The *cis*-acting elements required for insertion editing have been localized to a relatively small region in the vicinity of editing sites, but the details as to how editing sites are recognized and the identity of the *trans*-acting editing factor(s) required for insertion of these extra nucleotides remain to be elucidated. Two other mechanistically distinct forms of editing, 5′ tRNA editing and C-to-U conversion, have also been described, which proceed via two independent, posttranscriptional pathways. The relatively recent availability of genome and transcriptome sequence data has facilitated the identification of potential candidates for each of these activities and experiments to determine which of these factors are involved in the various forms of editing are underway.

J. Houtz · N. Cremona · J. M. Gott (✉)
Center for RNA Science and Therapeutics, 10900 Euclid Avenue, Case Western Reserve University, Cleveland, OH, USA
e-mail: jmg13@case.edu

J. Cruz-Reyes, M. Gray (eds.), *RNA Metabolism in Mitochondria*, Nucleic Acids and Molecular Biology 34, https://doi.org/10.1007/978-3-319-78190-7_8

8.1 Introduction

RNA processing and maturation of certain transcripts require site-specific changes in the nucleotide sequence or base composition that are not encoded at the DNA level. The process by which this maturation occurs is referred to as "RNA editing" and includes both nucleobase substitutions or conversions and nucleotide insertions and deletions. RNA editing occurs in all major classes of RNAs, including tRNAs, rRNAs, mRNAs, long noncoding RNAs (lncRNAs), and microRNAs (miRNAs), and is observed in diverse eukaryotic organisms including unicellular protists, worms, flies, plants, and mammals (reviewed in Knoop 2011). Depending on the context, RNA editing may alter either the folding stability of structural RNAs or the coding potential of mRNAs; these editing events often reveal cryptic genes, or "cryptogenes," whose transcripts are otherwise unrecognizable at the genome level (Gott et al. 2005; Sturm and Simpson 1990).

8.1.1 Internal Insertion/Deletion of Nucleotides

The phenomenon of RNA editing was first described in *Trypanosoma brucei* and *Crithidia fasciculata* by Benne and colleagues (Benne et al. 1986), who reported the insertion of four non-encoded uridines (Us) into the *cox2* mRNA. Further characterization of the kinetoplastid transcriptome revealed widespread, site-specific U insertions as well as U deletions (see Chap. 5). Shortly thereafter it was discovered that one or more extra nucleotides are added at a specific site in RNAs from various paramyxoviruses (Cattaneo et al. 1989; Thomas et al. 1988; Vidal et al. 1990) and throughout mitochondrial RNAs in the acellular slime mold *Physarum polycephalum* (Bundschuh et al. 2011; Mahendran et al. 1991).

8.1.2 Nucleotide Replacement at tRNA Ends

Transfer RNAs (tRNAs) have long been known to be subject to a vast array of internal nucleoside modifications, most of which cannot be encoded in the genome (Limbach et al. 1994; Söll 1971). The ends of mitochondrial tRNAs can also be targeted by editing events that correct mismatches within the acceptor stem. These changes occur at the nucleotide level and are effected via a variety of mechanisms. Nucleotide changes at the 5′ end of mitochondrial tRNAs were first reported in *Acanthamoeba castellanii* (Lonergan and Gray 1993) (see Chap. 7). Multiple examples of editing of the 3′ side of the acceptor stem were subsequently reported (reviewed in Betat et al. 2014), initially in tRNAs in snail mitochondria (Hatzoglou et al. 1995; Yokobori and Pääbo 1995).

8.1.3 Base Alterations

The term RNA editing was first defined as the insertion of non-encoded nucleotides, but this definition was expanded to include base changes upon the discovery of a single, site-specific C-to-U change within the cytoplasmic mammalian *apoB* mRNA (Powell et al. 1987). This was followed by the description of numerous C-to-U (Covello and Gray 1989; Gualberto et al. 1989; Hiesel et al. 1989), and less frequent U-to-C (Gualberto et al. 1990; Schuster et al. 1990), changes in mRNAs in plant mitochondria and, shortly thereafter, chloroplasts (Hoch et al. 1991) (see Chap. 9). The observation of a site-specific purine conversion within the glutamate receptor mRNA (Sommer et al. 1991) increased the repertoire of editing types further; the apparent A-to-G change at the cDNA level was attributed to the deamination of adenosine to inosine (A-to-I), an activity initially characterized by Bass and Weintraub (Bass and Weintraub 1988). Although A-to-I conversions in higher organisms have only been observed in RNAs of nuclear origin (Nishikura 2016), A-to-I changes have recently been reported in mitochondrial RNAs in diplonemids (Moreira et al. 2016).

8.1.4 Functions of Editing

RNA editing has a range of functions, some of which have yet to be uncovered (Gott and Emeson 2000; Knoop 2011). The insertion of nucleotides leads to the creation of open reading frames within mRNAs derived from cryptogenes that lack them. Base changes often result in codon alterations that can affect the sequence of the protein product. The creation of start and stop codons can occur via changes at either the base or nucleotide level. Likewise, splice sites and miRNA binding sites can be created or masked by editing, leading to the production of alternative proteins and/or changes in translatability and mRNA stability. Changes within noncoding RNAs generally affect secondary and tertiary interactions in structural RNAs such as tRNAs and rRNAs, while alterations within miRNAs and lncRNAs can affect interactions with their targets (Daniel et al. 2015; Nishikura 2016). The downstream effects of other editing events, such as those within RNAs derived from retro-elements, are unknown; these may result in functional changes, serve as molecular sponges, or simply be the result of off-target effects.

8.2 Editing in Mitochondria and Plastids

RNA editing is particularly prevalent in organelles. Widespread insertion editing is observed in the kinetoplasts of trypanosomes (see Chap. 5) and the mitochondria of *Physarum polycephalum* (this chapter) and *Diplonema papillatum* and related protists (Moreira et al. 2016). Mechanistic details are still lacking in many of these systems; however, based on their distinct characteristics, there are likely multiple editing mechanisms utilized in these disparate organisms.

Mitochondrial RNAs in diplonemids are also subject to base alterations, including both A-to-I and C-to-U changes (Moreira et al. 2016), and the interconversion of pyrimidines (C-to-U and U-to-C) is common in plant organelles (see Chap. 9). Mitochondrial mRNAs in higher plants are subject to hundreds of C-to-U changes, while plastid RNAs are generally edited at a lower level (Ichinose and Sugita 2016). Concomitant U-to-C changes are seen less frequently (Ichinose and Sugita 2016) but are common in chloroplast RNAs in some ferns (Wolf et al. 2004) and hornworts (Kugita et al. 2003; Yoshinaga et al. 1996) and in mitochondrial mRNAs in Placozoa (Burger et al. 2009; Steinhauser et al. 1999) and lycophytes (Grewe et al. 2011). A small number of C-to-U changes have also been reported in mitochondrial mRNAs in *Physarum* (Gott et al. 1993) and the protists *Naegleria* (Rüdinger et al. 2011) and *Acrasis* (Fu et al. 2014). While A-to-I and C-to-U changes in organelles have been attributed to deamination reactions, the mechanism underlying U-to-C changes has yet to be characterized in detail.

The mitochondria (Grewe et al. 2011; Lin et al. 2002; Shoguchi et al. 2015) and plastids (Mungpakdee et al. 2014; Zauner et al. 2004) of dinoflagellates and apicomplexans contain mRNAs that differ considerably from the sequence predicted based on their DNA sequences. These differences include pyrimidine (Py)-to-purine (Pu) and Pu-to-Py changes, as well as Py-Py and Pu-Pu exchanges. Most of these alterations cannot be ascribed to simple base interconversions and are thus likely to be the result of either base or nucleotide excision and replacement.

Mitochondrial tRNAs and rRNAs are also subject to a variety of editing events. Transfer RNA changes include replacement of nucleotides at the 5′ end in various Amoebozoa (Abad et al. 2014; Gott et al. 2010; Lonergan and Gray 1993) and chytridiomycete fungi (Bullerwell and Gray 2005) (see Chap. 7), addition of nucleotides at the 3′ end in metazoans (Betat et al. 2014), and internal nucleotide insertions in *Physarum* and *Didymium* (Antes et al. 1998). Mitochondrial rRNAs in myxomycetes contain over 100 internal nucleotide insertions (Table 8.1) (Bundschuh et al. 2011; Krishnan et al. 2007; Mahendran et al. 1994); less extensive editing of rRNAs has

Table 8.1 Summary of editing events in the mitochondria of *Physarum polycephalum*

Editing type	Total number	# of genes affected
+C	1255	46
+U	43	18
+G	2	2
+A	1	1
+AA	4	2
+UU	2	1
+UA	2	2
+UC/CU	9	5
+UG/GU	4	4
+CG/GC	2	2
−A	3	1
C → U	4	1
C → G	1	1
U → G	1	1

been reported in organelles from *Dictyostelium* (Barth et al. 1999), dinoflagellates (Dang and Green 2009; Zauner et al. 2004), and *Diplonema* (Moreira et al. 2016; Valach et al. 2014).

8.3 Editing in the Mitochondria of *Physarum polycephalum* and Related Myxomycetes

8.3.1 Characterization of Edited RNAs in Physarum polycephalum *Mitochondria*

Editing in *Physarum* mitochondria was initially discovered by Miller and colleagues, who reported the presence of 54 non-encoded C residues in the *Physarum atpA* mRNA (Mahendran et al. 1991). Characterization of additional mitochondrial RNAs led to the discovery of dinucleotide insertions (Gott et al. 1993; Mahendran et al. 1994), U insertions (Bundschuh et al. 2011; Gott et al. 1993; Mahendran et al. 1994), and four C-to-U changes (Gott et al. 1993), as well as the unusual finding that the rRNAs (Bullerwell et al. 2010; Krishnan et al. 2007; Mahendran et al. 1994) and four of the five tRNAs encoded in the mitochondria (Antes et al. 1998) also contain extra nucleotides. However, even after the entire *Physarum* mitochondrial genome was sequenced, many of the genes expected to be present could not be identified using standard gene-finding programs (Takano et al. 2001). The development of specialized algorithms (Beargie et al. 2008; Gott et al. 2005) led to the identification of additional genes; subsequent characterization of these mRNAs uncovered a site within the *nad2* mRNA where three encoded A residues are deleted (Gott et al. 2005). The entire complement of RNAs expressed in the mitochondria of plasmodial cells was ultimately identified by sequencing total mitochondrial RNA (Bundschuh et al. 2011). The RNA-seq data revealed 775 additional editing sites, including a single A insertion and 2 sites of single G insertion (Table 8.1). REDBASE, a searchable database of *Physarum* editing sites, has been established and is freely available (http://bioserv.mps.ohio-state.edu/redbase/) (Gott 2013).

The *Physarum* genome contains 45 cryptogenes, including the genes encoding 37 mRNAs, 5 tRNAs, and the large, small, and 5S-like rRNAs (Bundschuh et al. 2011). Added nucleotides typically make up ~4% of the nucleotides in mRNAs and ~2% in tRNAs and rRNAs. The overwhelming majority (94%) of the 1333 total editing events in *Physarum* mitochondria are sites of single cytidine (C) insertions. All but ten of these sites fall within the coding regions of mRNAs or within mature tRNAs and rRNAs. Nine of the ten extragenic edited sites fall within the 5′ or 3′ UTRs of mRNAs; only one C insertion is intergenic, falling between $tRNA^{Met2}$ and $tRNA^{Lys}$.

While RNA editing events only occur in specific contexts and/or developmental stages in some organisms, *Physarum* mitochondrial RNAs are edited in all life cycle stages (i.e., in both diploid and haploid forms) (Rundquist and Gott 1995). Editing in *Physarum* is highly accurate and exceptionally efficient, as incorrectly

inserted nucleotides and incompletely edited sites are undetectable in steady-state RNA pools in vivo (Byrne et al. 2002; Gott et al. 1993). The sole exception is the single intergenic C insertion site, which is only edited to ~50% (Bundschuh et al. 2011).

Editing patterns are similar between mRNAs, tRNAs, and rRNAs (Fig. 8.1). Excluding dinucleotide insertion sites, the minimum distance between editing sites is nine nucleotides. This distance constraint is likely to be related to the mechanism, as described below. There is a significant statistical bias in editing site contexts (Bundschuh et al. 2011; Miller et al. 1993); 59% of unambiguous C insertions (i.e., C residues not inserted next to an encoded C) follow a Pu-U, with A-U more commonly seen than G-U at positions -2 and -1, respectively. However, a large percentage of editing sites don't fit this pattern, and, given that the mitochondrial genome is ~74.1% AT, the significance of this bias is unclear. No consensus sequence is discernable, which is perhaps not surprising given the frequency of editing sites and the considerable constraints imposed by the need to conserve the amino acid sequence of essential mitochondrial proteins and the overall structure of noncoding RNAs (Fig. 8.1).

The first editing site within an mRNA is usually close to the initiation codon, and inserted nucleotides are present throughout the RNA, although regions upstream of the stop codon tend to have fewer editing sites (Gott 2013). Nucleotide insertions lead to frequent reading frameshifts (Fig. 8.1c), creating ORFs that encode proteins highly similar to those in mitochondria from other species. There is a definite codon bias, with 33% of unambiguous C insertion sites falling within the first position of a codon, 18% at the second position, and 49% in the third position (Bundschuh et al. 2011).

8.3.2 Editing in Other Myxomycetes

Mitochondrial RNAs in other myxomycetes are also edited to various extents. Analysis of the tRNAs encoded in the mitochondrial genomes of *Physarum* and the closely related *Didymium nigripes* showed similar patterns, although only one C insertion site is common to both species (Antes et al. 1998). A curious finding in these experiments was that while single tRNAs are fully edited, polycistronic tRNAs that are only partially edited can be detected in both species. Partially edited mRNAs are virtually undetectable, so it may be that incomplete editing in relatively stable tRNA precursors impairs tRNA processing.

To determine the extent of editing in more distantly related myxomycetes, Horton and Landweber (2000) examined editing patterns within a 1200 nucleotide region of the *cox1* mRNA from *Physarum polycephalum*, *Didymium nigripes*, *Stemonitis flavogenita*, *Arcyria cinerea,* and *Clastoderma debaryanum*. The extent of insertion editing in *Physarum*, *Didymium*, and *Stemonitis* is similar (34–40 C insertions, 1–4 U insertions, 3 dinucleotide insertions). The *cox1* mRNAs in the other two species are edited at a much lower level, with only

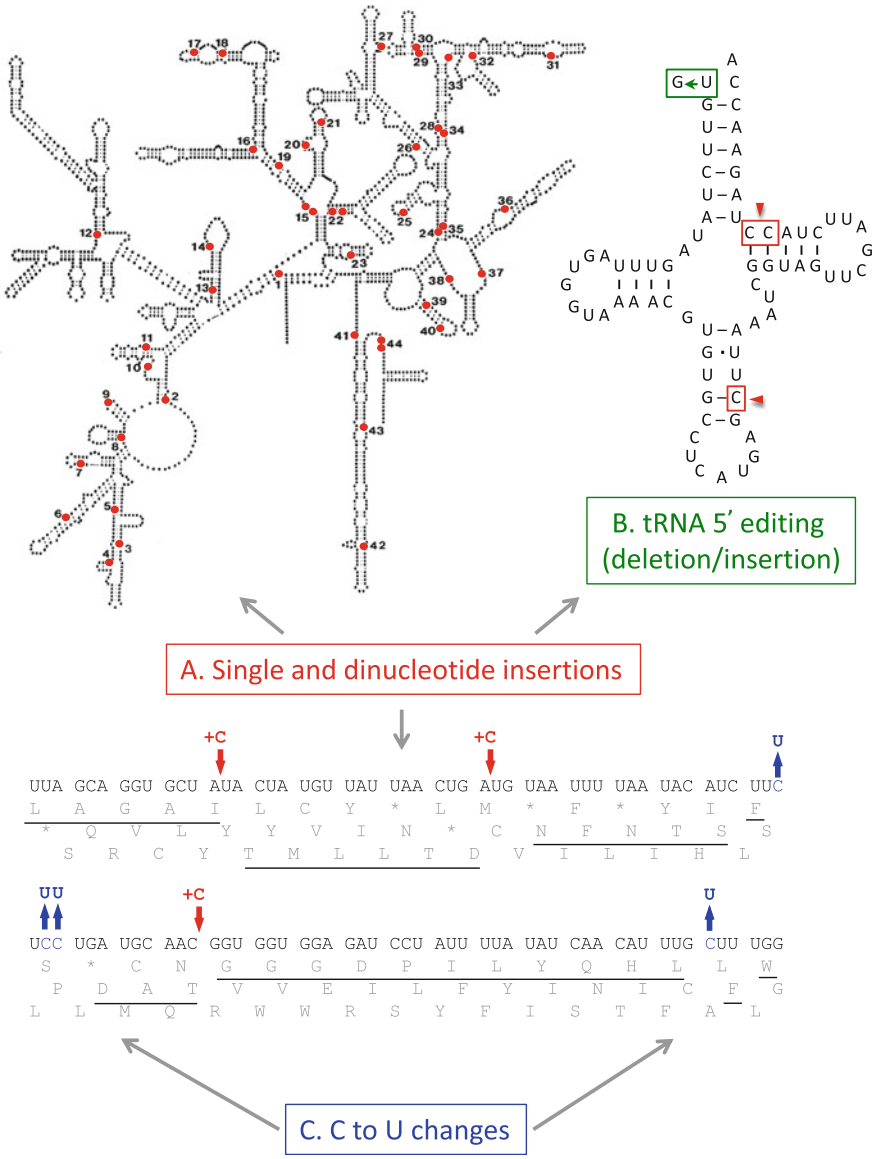

Fig. 8.1 *Physarum* mitochondrial RNAs are subject to three distinct forms of editing. (**a**) Non-encoded nucleotides are added co-transcriptionally to rRNAs (red dots, top left), tRNAs (red boxes, top right), and mRNAs (red downward arrows, bottom) as either single or dinucleotide insertions. (**b**) The nucleotide at the 5′ end of two of the tRNAs encoded in the mitochondrial genome is removed and replaced posttranscriptionally to complete the tRNA acceptor stem (green box, top right). (**c**) Four of the encoded C residues in the *cox1* mRNA are changed to U residues (blue letters, blue upward arrows, bottom) in a posttranscriptional process

4 U insertions in this region in *Clastoderma* and 4 U insertions and a single C added in *Arcyria*. Surprisingly, the pattern of C-to-U conversions was quite different; *Stemonitis* lacks C-to-U changes, while *Physarum* and *Didymium* each have four and *Arcyria* has two. Surprisingly, although all but 1 of these 10 changes fall in five codons within a 60 nucleotide span of the *cox1* mRNA, none is shared between species.

Silliker and colleagues have characterized the editing sites within 16 mRNAs encoded in the mitochondria of another myxomycete, *Didymium iridis* (Hendrickson and Silliker 2010; Traphagen et al. 2010). Overall patterns of C, U, and dinucleotide insertions are similar, but not identical to those in *Physarum polycephalum*, with sites of U and dinucleotide insertion more highly conserved than C insertion sites. There are also three C-to-U sites in these mRNAs, one in *cox1* and two in *cox2*, none of which is shared with *Physarum*. This work also described the first example of a single A insertion in myxomycete mitochondria (Hendrickson and Silliker 2010).

The most comprehensive analysis of myxomycete editing to date was carried out by Miller and colleagues (Krishnan et al. 2007), who looked at the distribution of editing sites within a conserved region of the mitochondrial small ribosomal rRNA (SSU) from seven different species. Alignment of mtDNA and cDNAs derived from *Physarum polycephalum*, *Physarum didermoides*, *Didymium nigripes*, *Didymium iridis*, *Lycogala epidendrum*, *Echinostelium minutum*, and *Stemonitis flavogenita* demonstrated that the pattern of editing sites is similar between species. In all species, editing sites are enriched in the most highly conserved regions of the rRNA. At least 27 of the 29 insertion sites fall within regions predicted to be base-paired and, based on covariation analyses, are involved in stabilization of common secondary structures. However, there was surprising variability in the location of inserted nucleotides. Of the 29 editing sites identified, none is conserved in all 7 species, and 14 insertion sites are present in only a single organism. The editing patterns in *Physarum polycephalum* and both species of *Didymium* are identical in this region, but none of these ten sites is shared by *Lycogala*. *Echinostelium* and *Lycogala* each utilize six unique insertion sites and share three others. Four of the sites in *Physarum didermoides* are in common with *Physarum polycephalum*, and three are shared with *Lycogala*, while one site is unique to this species. *Stemonitis* has six sites in common with *Physarum polycephalum* and one with *Echinostelium* but also has an unprecedented deletion of a single encoded U.

Despite the use of different editing sites, overall editing patterns are strikingly similar in terms of density and spacing (Horton and Landweber 2000; Krishnan et al. 2007; Bundschuh et al. 2011; Hendrickson and Silliker 2010; Traphagen et al. 2010). Each contains eight to ten editing sites within this region of the rRNA, and all sites within a given organism are at least nine nucleotides apart. Contexts are also similar, with editing sites being preceded by a Pu-U at a higher frequency than expected by chance. These similarities imply the existence of mechanistic constraints.

8.4 Editing Mechanisms in *Physarum polycephalum* Mitochondria

In organello and in vitro studies have led to the conclusion that there are at least three distinct editing mechanisms within *Physarum* mitochondria, comprising three types of editing events: (1) C-to-U substitutions in the cytochrome oxidase I (*coxI*) mRNA (Gott et al. 1993), (2) 5′ editing of $tRNA^{Met1}$ and $tRNA^{Met2}$ (Gott et al. 2010), and (3) mono- and dinucleotide insertions (Bundschuh et al. 2011) (Fig. 8.1). In stark contrast to editing mechanisms in most other systems, nucleotide insertion into *Physarum* mitochondrial RNAs occurs co-transcriptionally, with non-templated nucleotides incorporated at the 3′ end of nascent transcripts (Cheng et al. 2001). The mechanism underlying the AAA deletion within the *nad2* mRNA has not been examined directly, largely because this mRNA is expressed at very low levels in vitro (Gott et al. 2016). The complete absence of these three encoded nucleotides in steady-state RNA pools has been confirmed by primer extension sequencing of the mRNA (Gott et al. 2005), leading to the hypothesis that deletions are carried out co-transcriptionally, perhaps by the same apparatus as nucleotide insertions. As described below, the C-to-U substitutions and 5′ tRNA editing in *Physarum* mitochondria are posttranscriptional editing events.

8.4.1 tRNA Editing

The 5′ ends of two of the five mitochondrion-encoded tRNAs ($tRNA^{Met1}$ and $tRNA^{Met2}$) are subject to nucleotide replacement (Fig. 8.1b). These tRNAs were initially suggested to be edited in this manner by M. W. Gray (personal communication) based on the presence of predicted mismatches at the top of the acceptor stem, a pattern characteristic of edited tRNAs in *A. castellanii* mitochondria. This expectation was confirmed by in vitro labeling studies and sequencing of circularized tRNAs (Gott et al. 2010). The 5′-editing reaction displays striking similarities to the reaction catalyzed by $tRNA^{His}$ guanylyltransferase (Thg1), the enzyme responsible for addition of a single G to the 5′ end of $tRNA^{His}$ in yeast and other organisms (Gu et al. 2003). A search of available *Physarum* sequence data for Thg1-related genes resulted in the identification of two potential candidates for the tRNA editing activity. These *Physarum* sequences were then used to identify orthologs of these potential editing factors in the genomes of *A. castellanii*, *Dictyostelium discoideum*, and numerous other species (Jackman et al. 2012). As described in Chap. 7, biochemical characterization of the four Thg1-like proteins (TLPs) identified in *D. discoideum* confirmed that these proteins act on 5′ ends of tRNA (Abad et al. 2011) and identified DdiTLP3 as the enzyme responsible for 5′ tRNA editing in *D. discoideum* mitochondria (Long et al. 2016). Unlike many of the edited tRNAs in *A. castellanii* and *D. discoideum*, the *Physarum* tRNAs subject to

nucleotide replacement editing contain only a single mismatch at the top of the acceptor stem (Gott et al. 2010). This opens up the possibility that the properties of the *Physarum* enzyme(s) may differ slightly from the previously characterized enzymes.

8.4.2 C-to-U Editing

The determination that C-to-U editing in *Physarum* mitochondria is posttranscriptional is based on three lines of evidence: (1) the presence of C rather than U at these positions in otherwise edited RNAs made in vitro (Byrne and Gott 2002; Visomirski-Robic and Gott 1995), (2) S1 nuclease protection assays of labeled RNAs synthesized in isolated mitochondria (Visomirski-Robic and Gott 1995), and (3) sequencing of nascent (i.e., polymerase-associated) RNAs made in vivo, which contain a mixture of C and U at these sites (Byrne and Gott 2004). The four C-to-U changes in the *Physarum cox1* mRNA (Fig. 8.1c) are likely to be targeted by proteins similar to those present in plant mitochondria. Recognition of C-to-U sites in plant mitochondria requires PPR proteins, the majority of which contain a C-terminal DYW domain (see Chap. 9). Most organisms have few, if any, PPR proteins, and only a very small subset of sequenced genomes appears to encode PPR-DYW proteins (Schallenberg-Rüdinger et al. 2013). The presence of PPR-DYW proteins suggests these species may exhibit C-to-U editing, a prediction that has already been borne out in *Naegleria* (Rüdinger et al. 2011) and *Acrasis* (Fu et al. 2014). Approximately 100 PPR proteins are encoded in the *Physarum* genome, 16 of which contain a recognizable DYW domain (Schallenberg-Rüdinger et al. 2013). Five of these DYW domains are highly homologous to those present in plant editing factors, making them leading candidates for the *Physarum* C-to-U editing activity. Given that in its natural environment *Physarum* feeds on bacteria and decaying plant material, *Physarum* may have acquired the C-to-U editing activity via lateral gene transfer. Curiously, however, BLAST searches of *Physarum* genome and transcriptome data (Schaap et al. 2015) have failed to identify genes with significant homology to any of the auxiliary proteins shown to be involved in editing in plant mitochondria, suggesting that the *Physarum* C-to-U editing apparatus may be a simpler form of the editing machinery ("editosomes") found in plants (see Chap. 9). Alternatively, *Physarum* may encode auxiliary factors distinct from those in other organisms. Preliminary studies to identify and isolate factors required for C-to-U editing are underway (see below).

8.4.3 Insertion Editing

The pattern of nucleotide insertion first observed in *Physarum* mitochondrial RNAs is thus far unique to *Physarum* and closely related myxomycetes (Fig. 8.1a). Characterization of steady-state RNA pools indicated that essentially all RNAs were fully edited (Gott et al. 1993) and strategies implemented to look for unedited or partially

edited mRNAs failed to find potential intermediates. These findings led to the hypothesis that insertion editing was somehow linked to transcription. Based on the premise that editing occurs close to the site of RNA synthesis, initial efforts at developing an in vitro editing system focused on isolated organelles. In order to distinguish transcripts made in vitro from the vast pool of existing mitochondrial RNAs, run-on transcripts were labeled, and individual transcripts were isolated via S1 nuclease protection and analyzed via RNA fingerprinting (two-dimensional separation of RNA oligonucleotides via denaturing gel electrophoresis followed by thin-layer chromatography) (Visomirski-Robic and Gott 1995, 1997a, b). Transcripts synthesized in isolated mitochondria were efficiently edited under most conditions. However, under severely limiting CTP concentrations, run-on transcripts were largely unedited at C insertion sites, confirming that the mitochondrial genome is the editing template (Visomirski-Robic and Gott 1997a). Remarkably, upon restoration of editing conditions, the subsequently transcribed RNA was edited, while nucleotide insertion into previously synthesized, unedited RNA, was not observed. These data demonstrated that the editing machinery works in association with the transcriptional apparatus and that nucleotide insertion proceeds unidirectionally with a 5′ to 3′ polarity (Visomirski-Robic and Gott 1997a).

In order to further purify transcription/editing complexes, mitochondrial lysates were fractionated via gel filtration chromatography to enrich for native mitochondrial DNA (mtDNA) and its associated endogenous protein complement. These mtDNA-protein complexes, which support run-on transcription and editing, are referred to as "mitochondrial transcription elongation complexes" or mtTECs (Cheng and Gott 2000). Although these fractions remain editing competent, RNAs synthesized in vitro by these crude extracts are not fully edited (Cheng and Gott 2000), and a low level (~5%) of sites are mis-edited (Byrne et al. 2002). The two predominant forms of mis-editing are the insertion of a single G at C insertion sites and inter-site deletions (i.e., the precise deletion of encoded nucleotides between two editing sites). Importantly, all mis-editing events occur at bona fide insertion sites, suggesting that editing sites are uniquely demarcated and that editing site recognition and nucleotide insertion can be uncoupled (Byrne et al. 2002).

Editing-competent mtTECs were shown to lack detectable levels of nucleotide triphosphates (NTPs) (Cheng and Gott 2000), enabling the manipulation of the relative concentrations of each NTP during run-on transcription/editing. These experiments demonstrated that the level of C addition at individual insertion sites can be varied in a predictable pattern by lowering or raising the concentration of the templated NTP immediately downstream of the site relative to the CTP concentration in the reaction. High concentrations of the next encoded nucleotide significantly reduced the level of C insertion, while low levels of the following nucleotide greatly increased the extent of editing. These findings indicated that transcription and editing are competing processes, leading to the conclusion that non-encoded nucleotides are added to the 3′ end of nascent RNAs during transcription of the mitochondrial genome (Cheng et al. 2001) (Fig. 8.2).

The sequence elements that direct editing to specific sites have not been identified. Bioinformatics approaches have failed to uncover consensus sequences or conserved features within the mitochondrial DNA (Bundschuh et al. 2011; Chen et al. 2012).

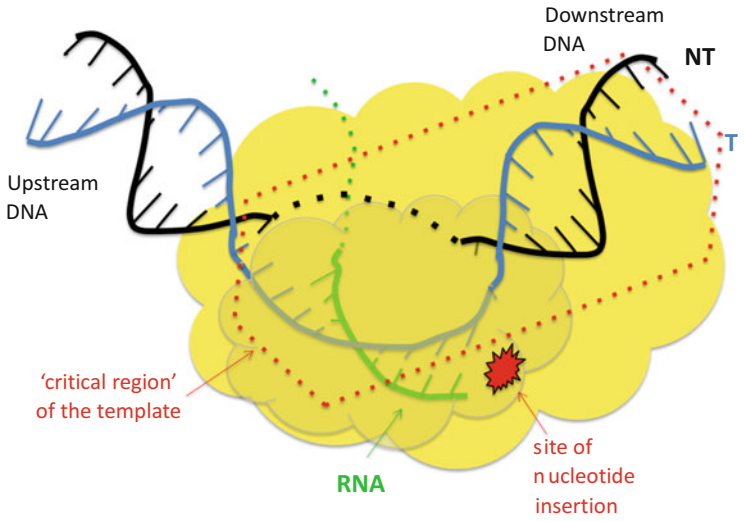

Fig. 8.2 Diagram depicting predicted regions of contact between the mitochondrial RNA polymerase and nucleic acids in the *Physarum* transcription/editing complex. The model is based on published structures of the bacteriophage T7 transcription elongation complex (Yin and Steitz 2002) in which the DNA upstream of the active site is unwound and contained in a 7–8 bp RNA-DNA hybrid and the downstream DNA is largely double-stranded. The *Physarum* mitochondrial RNA polymerase is depicted as yellow clouds, the DNA strands are shown in light blue (T, template strand) and black (NT, non-template strand), and the nascent RNA is shown in green. DNA and RNA regions predicted to be enveloped by the polymerase are shown as lighter colored lines beneath the darker portion of the polymerase. The site of non-templated nucleotide insertion is indicated by the red starburst. The region of the DNA template outlined with red dots has been demonstrated to be critical for nucleotide insertion. See text for details

Furthermore, the nature of the in vitro systems, consisting of a pool of mitochondrial genomes with preformed transcription complexes, precludes systematic mutagenesis of sequences surrounding editing sites. To circumvent these issues, "chimeric templates" were created by fusing mitochondrial sequences to DNA fragments associated with transcription/editing complexes (Fig. 8.3). This was accomplished by digestion of the mtTEC DNA with restriction endonucleases and ligation to either artificial DNA constructs (i.e., cloned *Physarum* genes or PCR-generated cassettes, Fig. 8.3a, b) or to other mtTEC fragments, leading to rearrangements of the native templates (Byrne and Gott 2002) (Fig. 8.3d). Run-on transcription from these templates revealed that editing is only supported when the native template is transcribed; no editing is observed during transcription of non-native DNA downstream of ligation junctions, even in instances where the DNA sequence is identical to that of the native gene (Byrne and Gott 2002). A second critical observation from chimeric template studies was that DNA fragments isolated from phenol-extracted mtTEC also fail to support editing when ligated to native mtTEC fragments (Byrne and Gott 2002) (Fig. 8.3c), indicating that any potential signals due to epigenetic modifications are insufficient to demarcate editing sites. In contrast, rearranged native mtTEC

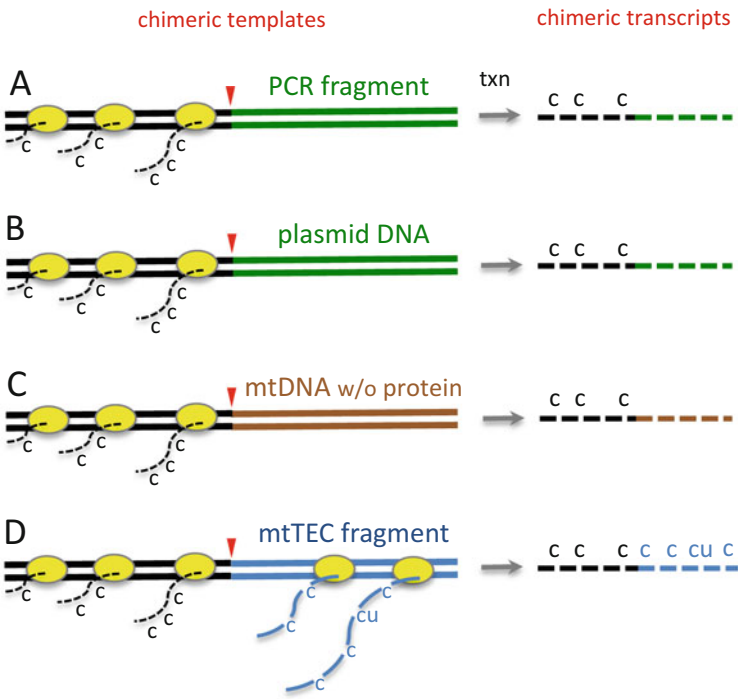

Fig. 8.3 Edited transcripts are only generated from the native portion of chimeric templates. Left: regions of the mitochondrial genome from native transcription elongation complexes (mtTECs) are depicted in black and blue, DNA fragments containing *Physarum* mitochondrial sequences generated by PCR or derived from plasmid DNA are shown in green, and DNA fragments isolated from phenol-extracted mtTECs are depicted in brown. Right: editing status of run-on transcripts derived from each type of chimeric template. See text for details

fragments, either in the form of linear or circularized fragments, maintain the same level of editing efficiency as intact mtTECs (Byrne and Gott 2002) (Fig. 8.3d). These results suggest a requirement for at least one template-associated *trans*-acting protein factor in the mechanism of *Physarum* insertion editing. Subsequent in vitro experiments showed that editing site recognition and insertion of non-encoded nucleotides are separable processes (Byrne et al. 2002). Curiously, single and dinucleotide insertion sites respond differently to changes in relative nucleotide concentrations in vitro, with patterns of editing and mis-editing differing even among dinucleotide sites, suggesting that additional factors may be needed for dinucleotide insertions (Byrne and Gott 2004).

To localize template sequences required for nucleotide insertion, restriction endonucleases that cut close to sites of editing were used to generate chimeric templates. Using this strategy it was possible to change template sequences either upstream or downstream of individual editing sites (Rhee et al. 2009) (Fig. 8.4). These experiments demonstrated that the *cis*-elements required for specifying C addition are limited to ~18 bp of DNA centered around the site of insertion, termed

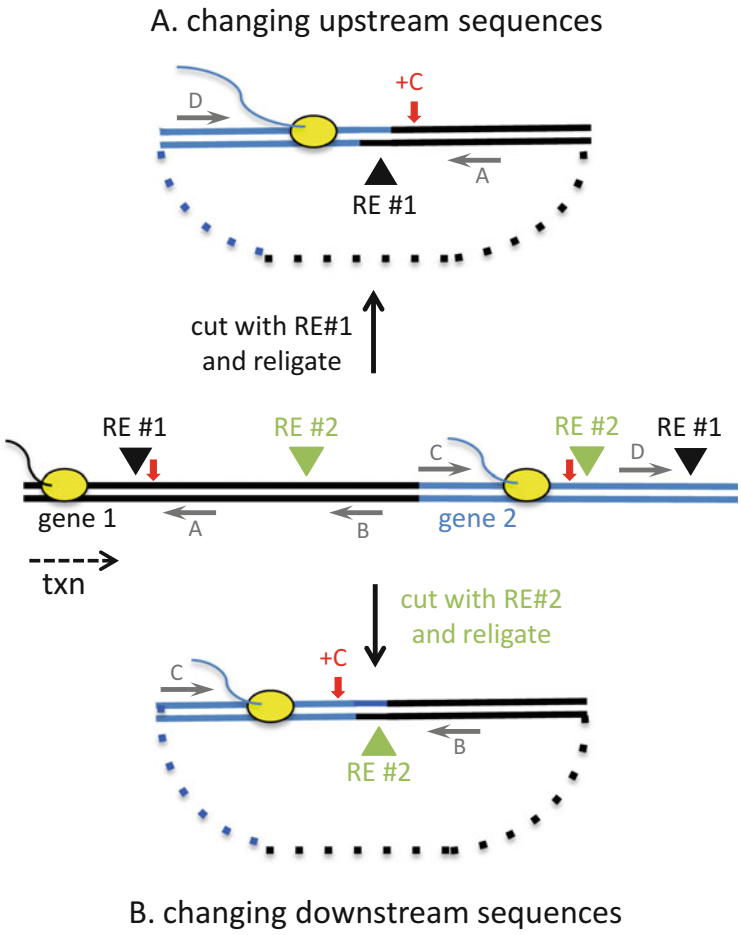

Fig. 8.4 Strategy for rearranging the native template to replace sequences proximal to editing sites. A restriction endonuclease (RE) with a recognition site close to a C insertion site (red arrows) is used to digest mitochondrial DNA in native mtTECs, and the resulting mixture of fragments is randomly ligated under conditions that favor intramolecular ligation. Run-on transcripts derived from these sub-genomic templates are characterized via RT-PCR using primers (gray arrows) that only yield a product from a specific chimeric template. (**a**) Strategy for changing sequences upstream of an editing site. (**b**) Strategy for changing sequences downstream of an editing site

the "critical region" of the template (Fig. 8.2). Template changes falling outside of this small window had no effect on editing at any of the sites tested (Rhee et al. 2009). Curiously, these experimentally defined limits are seemingly in conflict with the finding that no statistically significant nucleotide conservation patterns can be detected within this critical region (Bundschuh et al. 2011; Chen et al. 2012).

Because either upstream *or* downstream sequences, but not both, are changed in chimeric templates, it may be that redundant, longer range interactions on either side, are involved in "marking" an editing site (Chen et al. 2012). Alternatively, the editing apparatus may recognize subtle structural differences or utilize only a subset of nucleotides within the critical region at any one site (Chen et al. 2012; Gott 2013).

The sequence of the vast majority of RNAs transcribed from chimeric templates ("chimeric transcripts") contains all templated nucleotides near the junction. However, when template sequences very close to an editing site were altered, mis-editing events were occasionally observed; in all cases these mis-editing events occurred between the ligation junction on the template strand and the neighboring editing site (Rhee et al. 2009) (Fig. 8.5). Because both the restriction endonuclease and the DNA ligase used to generate the chimeric template are still present during the transcription/editing reactions, it is likely that in these rare cases, the template may have contained a nick at the ligation junction, potentially altering critical interactions with the transcription/editing complex. When upstream sequences within 9 bp of a site were changed, nucleotide insertions were not observed, but a small subset of products lacked the templated nucleotides between the site of ligation and the editing site immediately downstream of the ligation junction (Fig. 8.5a). Thus, the editing site appears to be recognized but not utilized in this context, suggesting that sequences downstream of an editing site, which are still intact in these chimeras, are involved in editing site recognition. When downstream sequences within 9–10 bp of an editing site were altered, the site was occasionally edited, but in these cases the edited products lacked the templated nucleotides between the insertion site and the ligation junction (Fig. 8.5b). These results, as well as the finding that changes in the upstream region appear to prevent nucleotide addition, implicate the template sequences 5′ of an editing site in nucleotide selection and/or insertion, and point to an additional role for the downstream region in templated extension from the unpaired nucleotide (Rhee et al. 2009).

The chimeric template experiments provided evidence that one or more *trans*-acting factors are required for insertion editing and suggested that such factors may be interacting with the mtDNA rather than traveling along the template in stable association with the mtRNAP. However, attempts to dissociate editing factors from mtTECs via salt washes or treatment with detergents have been unsuccessful, suggesting that either the mitochondrial RNA polymerase (mtRNAP) is capable of carrying out the editing reaction alone under these conditions or that, since editing can only be assayed in the context of run-on transcription, any factors required for editing are bound at least as tightly as the mtRNAP. Efforts to recapitulate editing using either native or recombinant mtRNAP failed to produce edited RNA. In both cases, only unedited RNA was synthesized, consistent with a need for additional editing factors. Current efforts aimed at identifying such factors are discussed below.

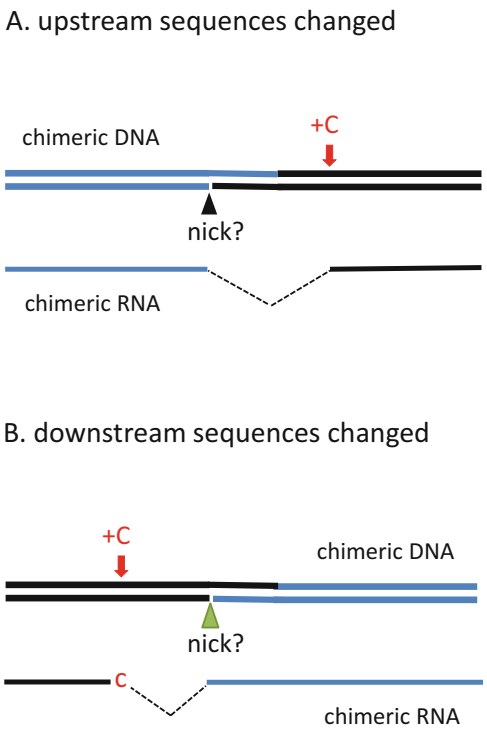

Fig. 8.5 Altering sequences within the critical region of chimeric templates produces distinct patterns of mis-editing. Templated nucleotides falling between the ligation junction on the template strand and the adjacent C insertion site (red arrows) were absent in a small subset of run-on transcripts derived from chimeric templates in which sequences within 9 bp of the editing site were altered (Rhee et al. 2009). (**a**) Chimeric templates in which sequences upstream of a C insertion site were changed occasionally yielded small deletions that lacked the C insertion. (**b**) Chimeric templates in which sequences downstream of a C insertion site were changed resulted in small deletions that often contained the non-encoded C. In each case, editing at sites outside of the critical region was not affected

8.5 Insertion Editing Models

The finding that insertion editing is closely linked to transcription places a number of constraints on potential editing mechanisms. One model considered initially was that transcription, editing, and translation are all tightly coupled to ensure that only edited RNAs are translated. However, the finding that rRNAs and tRNAs display similar editing patterns makes an essential link to translation doubtful. In viral systems (Kolakofsky 2016; Volchkov et al. 1995) nucleotide insertion generally occurs within a homopolymeric tract ("slippery site") within the RNA template, but the context of *Physarum* insertion sites is quite different, and only about a third of the inserted Cs are adjacent to even a single encoded C (Bundschuh et al. 2011). There are also significant mechanistic differences between the two systems in that varying

nucleotide concentrations in *Physarum* in vitro systems display a pattern opposite to that observed during "stuttering" by the viral polymerase within the homopolymer tract (Cheng et al. 2001; Visomirski-Robic and Gott 1997a).

8.5.1 Cis-*Acting Elements*

The *cis*-acting signals that direct nucleotide insertion into *Physarum* mitochondrial RNAs could potentially be located in the mtDNA, the RNA transcript, or both (Fig. 8.2). Based on the results with chimeric DNAs, the template elements required for specifying C addition are limited to ~9 base pairs (bp) upstream of a C insertion site and 9–10 bp downstream of the site (Rhee et al. 2009). Unfortunately, the currently available in vitro systems do not allow testing of individual strands of the DNA template (i.e., template vs. non-template strand) and do not distinguish whether the information resides at the level of DNA or RNA.

The finding that editing sites are spaced at least nine nucleotides apart is particularly intriguing given the boundaries of the essential template elements and may be indicative of physical or mechanistic constraints within the transcription/editing complex (Fig. 8.2). There is extensive homology between the polymerase domains of single-subunit mitochondrial RNA polymerases, including the *Physarum* mitochondrial RNA polymerase, and the closely related RNA polymerases from bacteriophages T7, T3, and SP6 (Miller et al. 2006; Gott and Rhee 2008). In high-resolution structures of T7 RNA polymerase transcription elongation complexes (Tahirov et al. 2002; Yin and Steitz 2002), the enzyme is associated with ~10 bp of the DNA upstream of the catalytic site and ~10 bp of downstream DNA. The DNA downstream of the active site is wrapped around the surface of the polymerase and is largely double-stranded. In contrast, the DNA upstream of the site of nucleotide addition is unwound. This portion of the template strand is largely base-paired with the nascent RNA, forming a ~7–8 bp RNA/DNA hybrid enveloped by the polymerase. The non-template DNA strand makes sequence-independent contacts with the outer surface of the polymerase, reforming a duplex as the template strand emerges from the enzyme. Thus, it is extremely likely that the regions of the DNA template required for insertion editing are very close to, or in direct contact with, the *Physarum* mtRNAP (Fig. 8.2).

In chimeric templates that change the sequence of the DNA upstream of an editing site, the primary sequence of the upstream RNA is also changed. Given that all but 9 bp of the upstream template are dispensable for editing, any requirement for specific RNA sequences 5′ of an insertion site is therefore limited to the nine nucleotides immediately upstream of that site. Because ~12 nucleotides of RNA are protected from nuclease digestion in T7 RNA polymerase elongation complexes (Huang and Sousa 2000), the region of the transcript that could influence editing is expected to be within the confines of the *Physarum* mtRNA polymerase, with nearly all of this region engaged in the RNA-DNA hybrid. This model is supported by in vitro experiments involving the removal of upstream RNA via oligonucleotide-directed RNase H cleavage prior to run-on synthesis, which demonstrated that

upstream RNA having just emerged from the elongation complex is not required for editing (Majewski and Gott, unpublished data). Because non-encoded nucleotides are added to the 3′ end of nascent transcripts (Cheng et al. 2001), involvement of RNA sequences downstream of an editing site would not be expected. While it cannot be formally excluded that the *Physarum* polymerase "backtracks" to insert a non-encoded nucleotide after incorporating the templated nucleotide(s) downstream, both the RNase H experiment and the nucleotide concentration effects mentioned above make this scenario unlikely.

8.5.2 Trans-acting Factors

Taken together, the results from both the chimeric template experiments and experiments with purified native and recombinant mtRNAP argue for the existence of one or more *trans*-acting editing factors. In vitro experiments also suggest that dinucleotide insertions may require additional factors (Byrne and Gott 2004). Such factors could conceivably specify the location of editing sites guided by either site-specific DNA sequence or structure. However, there is currently no evidence to support the presence of proteins bound in the vicinity of editing sites. The mitochondrial DNA in mtTECs appears to be fully accessible to restriction enzymes, including those whose recognition sites overlap editing sites (Byrne 2004; Rhee et al. 2009). Likewise, DNA foot-printing experiments failed to identify protected regions within the DNA template, arguing against (but not precluding) the static binding of editing factors at editing sites. It is possible that editing factors are not tightly bound or exchange rapidly with transcription complex elements, although, these possibilities are seemingly inconsistent with the findings that transcription/editing complexes can be purified via gel filtration chromatography (Cheng and Gott 2000) and affinity selection (Houtz, unpublished) and are stable in the presence of salts and detergents (Rhee and Gott, unpublished).

It is unknown whether it is the *Physarum* mtRNAP and/or auxiliary factors that catalyze the actual insertion of non-encoded nucleotides. However, the polymerase must be intimately associated with the editing reaction based on its involvement in the events preceding and following the editing reaction: (1) the mtRNAP must halt (i.e., pause) transcription once it reaches an editing site long enough to allow for both selection and insertion of the non-templated nucleotide(s) and (2) the nascent RNA must be extended in a template-directed fashion (Cheng et al. 2001). To accomplish this, the *Physarum* mtRNAP must be capable of accommodating (an) "extra," unpaired nucleotide(s) at the growing end of the RNA chain as well as a RNA-DNA hybrid containing one to two unpaired nucleotides. It is possible, therefore, that the need for additional *trans*-acting factors may reflect the need to facilitate and/or stabilize transitions between transcription and editing conformations in the polymerase itself.

8.6 Current Work and Future Directions

The identification and subsequent characterization of *trans*-acting editing factors are of paramount importance to the establishment of reconstituted in vitro editing systems and, ultimately, the elucidation of *Physarum* editing mechanisms. Using a combination of biochemical and bioinformatics approaches, candidate factors for each form of editing have been identified, and their potential roles are now being explored.

8.6.1 Factors Involved in Insertion Editing

A top priority of current work is the identification of factors involved in insertion editing. Although the isolation of mtTECs via gel filtration removes ~95% of the proteins present in the mitochondrial lysate (Cheng and Gott 2000), these fractions still contain a complex mixture of proteins, most of which are unlikely to be involved in editing. To reduce the complement of proteins in these fractions further, an affinity selection approach involving the isolation of defined portions of the mitochondrial genome along with their associated proteins has been developed. Importantly, these affinity-selected complexes retain editing activity and contain a limited number of proteins. All proteins co-purifying with a subset of mitochondrial DNA fragments encoding edited RNAs were analyzed via mass spectrometry, and each peptide was mapped to the reference *Physarum* transcriptome (Schaap et al. 2015). The resulting dataset contains proteins with no significant homology to known or predicted proteins as well as proteins predicted to contain nucleic acid-binding domains. These high-priority candidates are being targeted in ongoing gene silencing experiments to determine which of these are essential for nucleotide insertion. In addition, since it is highly likely that any *trans*-acting factor involved in insertion editing interacts at some point with the *Physarum* mtRNAP, yeast two-hybrid screens are being used to assess whether any of these candidates physically interact with the polymerase.

8.6.2 Factors Involved in C-to-U Editing

Until recently, the mechanism underlying the C-to-U changes in *Physarum* mitochondria has been largely unexplored. The existence of *Physarum* genes encoding PPR-DYW proteins suggests that the machinery responsible for C-to-U changes in *Physarum* may be similar to that in plant mitochondria (see Chap. 9). However, the apparent lack of *Physarum* orthologs of most of the general plant editing factors raises the possibility that *Physarum* C-to-U editosomes may differ significantly from those found in plants. A number of independent methods are currently being used to identify and isolate factors that specifically bind to and/or act upon RNAs containing

one or more C-to-U sites. Initial gel shift experiments provided evidence that one or more factors capable of specifically binding to C-to-U substrate RNAs are present in mitochondrial lysates. This finding was supported by the detection of proteins that physically cross-link to these RNAs upon UV irradiation. Affinity selection of these proteins is underway, with a goal of identifying initial targets for gene silencing. These experiments will lay the groundwork for the establishment of in vitro editing assays and the ultimate identification of the full complement of proteins required for C-to-U conversion in *Physarum* mitochondria.

8.7 Concluding Remarks and Outstanding Questions

There are at least three distinct editing mechanisms at play within *Physarum* mitochondria. Multiple examples of C-to-U and 5′ tRNA editing are found in nature, and there are commonalities between these forms of editing and those observed in myxomycetes. Evidence suggests, however, that features of these forms of editing may be idiosyncratic in *Physarum*. Although significant progress has been made toward deciphering the unique mechanism by which non-encoded nucleotides are precisely added to create functional RNAs in *Physarum* mitochondria, many questions remain. These fall into three broad categories: the source of the information (*How are the sites of editing and the identity of the nucleotide to be added specified? Are separate signals required?*), the composition of the editing apparatus (*What trans-acting factors are required? Does the mtRNAP participate in the editing reaction?*), and the actual mechanics of the reaction (*What role does each component play? What signals the mtRNAP to pause transcription at an editing site and to resume once the site is edited? Is there a template for nucleotide insertion? If so, is it composed of nucleic acids, protein, or both? Are additional cofactors required? What enzyme is responsible for adding the non-encoded nucleotides? How does the polymerase extend the RNA chain from an unpaired nucleotide in a template-directed manner? How is (are) the extra nucleotide(s) accommodated within the RNA/DNA hybrid? What is the basis for the extraordinary efficiency of the editing reaction? How is the accuracy achieved? Is there a proofreading function?*). Clearly, the identification of *trans*-acting factors will be a key step in beginning to dissect this process mechanistically.

References

Abad MG, Long Y, Willcox A, Gott JM, Gray MW, Jackman JE (2011) A role for $tRNA^{His}$ guanylyltransferase (Thg1)-like proteins from *Dictyostelium discoideum* in mitochondrial 5′-tRNA editing. RNA 17(4):613–623. https://doi.org/10.1261/rna.2517111

Abad MG, Long Y, Kinchen RD, Schindel ET, Gray MW, Jackman JE (2014) Mitochondrial tRNA 5′-editing in *Dictyostelium discoideum* and *Polysphondylium pallidum*. J Biol Chem 289(22): 15155–15165. https://doi.org/10.1074/jbc.M114.561514

Antes T, Costandy H, Mahendran R, Spottswood M, Miller D (1998) Insertional editing of mitochondrial tRNAs of *Physarum polycephalum* and *Didymium nigripes*. Mol Cell Biol 18(12):7521–7527

Barth C, Greferath U, Kotsifas M, Fisher PR (1999) Polycistronic transcription and editing of the mitochondrial small subunit (SSU) ribosomal RNA in *Dictyostelium discoideum*. Curr Genet 36(1–2):55–61

Bass BL, Weintraub H (1988) An unwinding activity that covalently modifies its double-stranded RNA substrate. Cell 55(6):1089–1098

Beargie C, Liu T, Corriveau M, Lee HY, Gott J, Bundschuh R (2008) Genome annotation in the presence of insertional RNA editing. Bioinformatics 24(22):2571–2578

Benne R, Van den Burg J, Brakenhoff JP, Sloof P, Van Boom JH, Tromp MC (1986) Major transcript of the frameshifted *coxII* gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. Cell 46(6):819–826

Betat H, Long Y, Jackman JE, Mörl M (2014) From end to end: tRNA editing at 5′- and 3′-terminal positions. Int J Mol Sci 15(12):23975–23998. https://doi.org/10.3390/ijms151223975

Bullerwell CE, Gray MW (2005) *In vitro* characterization of a tRNA editing activity in the mitochondria of *Spizellomyces punctatus*, a chytridiomycete fungus. J Biol Chem 280(4): 2463–2470. https://doi.org/10.1074/jbc.M411273200

Bullerwell CE, Burger G, Gott JM, Kourennaia O, Schnare MN, Gray MW (2010) Abundant 5S rRNA-like transcripts encoded by the mitochondrial genome in amoebozoa. Eukaryot Cell 9(5):762–773

Bundschuh R, Altmüller J, Becker C, Nürnberg P, Gott JM (2011) Complete characterization of the edited transcriptome of the mitochondrion of *Physarum polycephalum* using deep sequencing of RNA. Nucleic Acids Res 39(14):6044–6055

Burger G, Yan Y, Javadi P, Lang BF (2009) Group I-intron *trans*-splicing and mRNA editing in the mitochondria of placozoan animals. Trends Genet 25(9):381–386. https://doi.org/10.1016/j.tig.2009.07.003

Byrne EM (2004) Chimeric templates and assays used to study *Physarum* cotranscriptional insertional editing *in vitro*. Methods Mol Biol 265:293–314

Byrne EM, Gott JM (2002) Cotranscriptional editing of *Physarum* mitochondrial RNA requires local features of the native template. RNA 8(9):1174–1185

Byrne EM, Gott JM (2004) Unexpectedly complex editing patterns at dinucleotide insertion sites in *Physarum* mitochondria. Mol Cell Biol 24(18):7821–7828

Byrne EM, Stout A, Gott JM (2002) Editing site recognition and nucleotide insertion are separable processes in *Physarum* mitochondria. EMBO J 21(22):6154–6161

Cattaneo R, Kaelin K, Baczko K, Billeter MA (1989) Measles virus editing provides an additional cysteine-rich protein. Cell 56(5):759–764

Chen C, Frankhouser D, Bundschuh R (2012) Comparison of insertional RNA editing in myxomycetes. PLoS Comput Biol 8(2):e1002400. https://doi.org/10.1371/journal.pcbi.1002400

Cheng YW, Gott JM (2000) Transcription and RNA editing in a soluble *in vitro* system from *Physarum* mitochondria. Nucleic Acids Res 28(19):3695–3701

Cheng YW, Visomirski-Robic LM, Gott JM (2001) Non-templated addition of nucleotides to the 3′ end of nascent RNA during RNA editing in *Physarum*. EMBO J 20(6):1405–1414. https://doi.org/10.1093/emboj/20.6.1405

Covello PS, Gray MW (1989) RNA editing in plant mitochondria. Nature 341(6243):662–666. https://doi.org/10.1038/341662a0

Dang Y, Green BR (2009) Substitutional editing of *Heterocapsa triquetra* chloroplast transcripts and a folding model for its divergent chloroplast 16S rRNA. Gene 442(1–2):73–80. https://doi.org/10.1016/j.gene.2009.04.006

Daniel C, Lagergren J, Ohman M (2015) RNA editing of non-coding RNA and its role in gene regulation. Biochimie 117:22–27. https://doi.org/10.1016/j.biochi.2015.05.020

Fu CJ, Sheikh S, Miao W, Andersson SG, Baldauf SL (2014) Missing genes, multiple ORFs, and C-to-U type RNA editing in *Acrasis kona* (Heterolobosea, Excavata) mitochondrial DNA. Genome Biol Evol 6(9):2240–2257. https://doi.org/10.1093/gbe/evu180

Gott JM (2013) Mechanisms and functions of RNA editing in *Physarum polycephalum*. In: Mass S (ed) RNA editing: current research and future trends. Caister Academic Press, Norfolk, pp 17–40

Gott JM, Emeson RB (2000) Functions and mechanisms of RNA editing. Annu Rev Genet 34: 499–531. https://doi.org/10.1146/annurev.genet.34.1.499

Gott JM, Rhee AC (2008) Insertion/deletion editing in *Physarum polycephalum*. In: Goringer HU (ed) RNA editing. Springer, Berlin, pp 85–104

Gott JM, Visomirski LM, Hunter JL (1993) Substitutional and insertional RNA editing of the cytochrome c oxidase subunit 1 mRNA of *Physarum polycephalum*. J Biol Chem 268(34): 25483–25486

Gott JM, Parimi N, Bundschuh R (2005) Discovery of new genes and deletion editing in *Physarum* mitochondria enabled by a novel algorithm for finding edited mRNAs. Nucleic Acids Res 33(16):5063–5072. https://doi.org/10.1093/nar/gki820

Gott JM, Somerlot BH, Gray MW (2010) Two forms of RNA editing are required for tRNA maturation in *Physarum* mitochondria. RNA 16(3):482–488. https://doi.org/10.1261/rna.1958810

Gott JM, Naegele GM, Howell SJ (2016) Electroporation of DNA into *Physarum polycephalum* mitochondria: effects on transcription and RNA editing in isolated organelles. Genes (Basel) 7(12). https://doi.org/10.3390/genes7120128

Grewe F, Herres S, Viehöver P, Polsakiewicz M, Weisshaar B, Knoop V (2011) A unique transcriptome: 1782 positions of RNA editing alter 1406 codon identities in mitochondrial mRNAs of the lycophyte *Isoetes engelmannii*. Nucleic Acids Res 39(7):2890–2902. https://doi.org/10.1093/nar/gkq1227

Gu W, Jackman JE, Lohan AJ, Gray MW, Phizicky EM (2003) $tRNA^{His}$ maturation: an essential yeast protein catalyzes addition of a guanine nucleotide to the 5′ end of $tRNA^{His}$. Genes Dev 17(23):2889–2901. https://doi.org/10.1101/gad.1148603

Gualberto JM, Lamattina L, Bonnard G, Weil JH, Grienenberger JM (1989) RNA editing in wheat mitochondria results in the conservation of protein sequences. Nature 341(6243):660–662. https://doi.org/10.1038/341660a0

Gualberto JM, Weil JH, Grienenberger JM (1990) Editing of the wheat *coxIII* transcript: evidence for twelve C to U and one U to C conversions and for sequence similarities around editing sites. Nucleic Acids Res 18(13):3771–3776

Hatzoglou E, Rodakis GC, Lecanidou R (1995) Complete sequence and gene organization of the mitochondrial genome of the land snail *Albinaria coerulea*. Genetics 140(4):1353–1366

Hendrickson PG, Silliker ME (2010) RNA editing in six mitochondrial ribosomal protein genes of *Didymium iridis*. Curr Genet 56(3):203–213. https://doi.org/10.1007/s00294-010-0292-4

Hiesel R, Wissinger B, Schuster W, Brennicke A (1989) RNA editing in plant mitochondria. Science 246(4937):1632–1634

Hoch B, Maier RM, Appel K, Igloi GL, Kössel H (1991) Editing of a chloroplast mRNA by creation of an initiation codon. Nature 353(6340):178–180. https://doi.org/10.1038/353178a0

Horton TL, Landweber LF (2000) Evolution of four types of RNA editing in myxomycetes. RNA 6(10):1339–1346

Huang J, Sousa R (2000) T7 RNA polymerase elongation complex structure and movement. J Mol Biol 303(3):347–358. https://doi.org/10.1006/jmbi.2000.4150

Ichinose M, Sugita M (2016) RNA editing and its molecular mechanism in plant organelles. Genes (Basel) 8(1). https://doi.org/10.3390/genes8010005

Jackman JE, Gott JM, Gray MW (2012) Doing it in reverse: 3′-to-5′ polymerization by the Thg1 superfamily. RNA 18(5):886–899. https://doi.org/10.1261/rna.032300.112

Knoop V (2011) When you can't trust the DNA: RNA editing changes transcript sequences. Cell Mol Life Sci 68(4):567–586. https://doi.org/10.1007/s00018-010-0538-9

Kolakofsky D (2016) Paramyxovirus RNA synthesis, mRNA editing, and genome hexamer phase: a review. Virology 498:94–98. https://doi.org/10.1016/j.virol.2016.08.018

Krishnan U, Barsamian A, Miller DL (2007) Evolution of RNA editing sites in the mitochondrial small subunit rRNA of the Myxomycota. Methods Enzymol 424:197–220. https://doi.org/10.1016/S0076-6879(07)24009-1

Kugita M, Yamamoto Y, Fujikawa T, Matsumoto T, Yoshinaga K (2003) RNA editing in hornwort chloroplasts makes more than half the genes functional. Nucleic Acids Res 31(9):2417–2423

Limbach PA, Crain PF, McCloskey JA (1994) Summary: the modified nucleosides of RNA. Nucleic Acids Res 22(12):2183–2196

Lin S, Zhang H, Spencer DF, Norman JE, Gray MW (2002) Widespread and extensive editing of mitochondrial mRNAs in dinoflagellates. J Mol Biol 320(4):727–739

Lonergan KM, Gray MW (1993) Editing of transfer RNAs in *Acanthamoeba castellanii* mitochondria. Science 259(5096):812–816

Long Y, Abad MG, Olson ED, Carrillo EY, Jackman JE (2016) Identification of distinct biological functions for four 3′-5′ RNA polymerases. Nucleic Acids Res 44(17):8395–8406. https://doi.org/10.1093/nar/gkw681

Mahendran R, Spottswood MR, Miller DL (1991) RNA editing by cytidine insertion in mitochondria of *Physarum polycephalum*. Nature 349(6308):434–438. https://doi.org/10.1038/349434a0

Mahendran R, Spottswood MS, Ghate A, Ling ML, Jeng K, Miller DL (1994) Editing of the mitochondrial small subunit rRNA in *Physarum polycephalum*. EMBO J 13(1):232–240

Miller D, Mahendran R, Spottswood M, Costandy H, Wang S, Ling ML, Yang N (1993) Insertional editing in mitochondria of *Physarum*. Semin Cell Biol 4(4):261–266

Miller ML, Antes TJ, Qian F, Miller DL (2006) Identification of a putative mitochondrial RNA polymerase from *Physarum polycephalum*: characterization, expression, purification, and transcription *in vitro*. Curr Genet 49(4):259–271. https://doi.org/10.1007/s00294-005-0053-y

Moreira S, Valach M, Aoulad-Aissa M, Otto C, Burger G (2016) Novel modes of RNA editing in mitochondria. Nucleic Acids Res 44(10):4907–4919. https://doi.org/10.1093/nar/gkw188

Mungpakdee S, Shinzato C, Takeuchi T, Kawashima T, Koyanagi R, Hisata K, Tanaka M, Goto H, Fujie M, Lin S, Satoh N, Shoguchi E (2014) Massive gene transfer and extensive RNA editing of a symbiotic dinoflagellate plastid genome. Genome Biol Evol 6(6):1408–1422. https://doi.org/10.1093/gbe/evu109

Nishikura K (2016) A-to-I editing of coding and non-coding RNAs by ADARs. Nat Rev Mol Cell Biol 17(2):83–96. https://doi.org/10.1038/nrm.2015.4

Powell LM, Wallis SC, Pease RJ, Edwards YH, Knott TJ, Scott J (1987) A novel form of tissue-specific RNA processing produces apolipoprotein-B48 in intestine. Cell 50(6):831–840

Rhee AC, Somerlot BH, Parimi N, Gott JM (2009) Distinct roles for sequences upstream of and downstream from *Physarum* editing sites. RNA 15(9):1753–1765. https://doi.org/10.1261/rna.1668309

Rüdinger M, Fritz-Laylin L, Polsakiewicz M, Knoop V (2011) Plant-type mitochondrial RNA editing in the protist *Naegleria gruberi*. RNA 17(12):2058–2062. https://doi.org/10.1261/rna.02962911

Rundquist BA, Gott JM (1995) RNA editing of the *coI* mRNA throughout the life cycle of *Physarum polycephalum*. Mol Gen Genet 247(3):306–311

Schaap P, Barrantes I, Minx P, Sasaki N, Anderson RW, Benard M, Biggar KK, Buchler NE, Bundschuh R, Chen X, Fronick C, Fulton L, Golderer G, Jahn N, Knoop V, Landweber LF, Maric C, Miller D, Noegel AA, Peace R, Pierron G, Sasaki T, Schallenberg-Rüdinger M, Schleicher M, Singh R, Spaller T, Storey KB, Suzuki T, Tomlinson C, Tyson JJ, Warren WC, Werner ER, Werner-Felmayer G, Wilson RK, Winckler T, Gott JM, Glöckner G, Marwan W (2015) The *Physarum polycephalum* genome reveals extensive use of prokaryotic two-component and metazoan-type tyrosine kinase signaling. Genome Biol Evol 8(1):109–125. https://doi.org/10.1093/gbe/evv237

Schallenberg-Rüdinger M, Lenz H, Polsakiewicz M, Gott JM, Knoop V (2013) A survey of PPR proteins identifies DYW domains like those of land plant RNA editing factors in diverse eukaryotes. RNA Biol 10(9):1549–1556. https://doi.org/10.4161/rna.25755

Schuster W, Hiesel R, Wissinger B, Brennicke A (1990) RNA editing in the cytochrome *b* locus of the higher plant *Oenothera berteriana* includes a U-to-C transition. Mol Cell Biol 10(5): 2428–2431

Shoguchi E, Shinzato C, Hisata K, Satoh N, Mungpakdee S (2015) The large mitochondrial genome of *Symbiodinium minutum* reveals conserved noncoding sequences between dinoflagellates and apicomplexans. Genome Biol Evol 7(8):2237–2244. https://doi.org/10.1093/gbe/evv137

Söll D (1971) Enzymatic modification of transfer RNA. Science 173(3994):293–299

Sommer B, Kohler M, Sprengel R, Seeburg PH (1991) RNA editing in brain controls a determinant of ion flow in glutamate-gated channels. Cell 67(1):11–19

Steinhauser S, Beckert S, Capesius I, Malek O, Knoop V (1999) Plant mitochondrial RNA editing. J Mol Evol 48(3):303–312

Sturm NR, Simpson L (1990) Kinetoplast DNA minicircles encode guide RNAs for editing of cytochrome oxidase subunit III mRNA. Cell 61(5):879–884

Tahirov TH, Temiakov D, Anikin M, Patlan V, McAllister WT, Vassylyev DG, Yokoyama S (2002) Structure of a T7 RNA polymerase elongation complex at 2.9 Å resolution. Nature 420(6911):43–50. https://doi.org/10.1038/nature01129

Takano H, Abe T, Sakurai R, Moriyama Y, Miyazawa Y, Nozaki H, Kawano S, Sasaki N, Kuroiwa T (2001) The complete DNA sequence of the mitochondrial genome of *Physarum polycephalum*. Mol Gen Genet 264(5):539–545

Thomas SM, Lamb RA, Paterson RG (1988) Two mRNAs that differ by two nontemplated nucleotides encode the amino coterminal proteins P and V of the paramyxovirus SV5. Cell 54(6): 891–902

Traphagen SJ, Dimarco MJ, Silliker ME (2010) RNA editing of 10 *Didymium iridis* mitochondrial genes and comparison with the homologous genes in *Physarum polycephalum*. RNA 16(4): 828–838. https://doi.org/10.1261/rna.1989310

Valach M, Moreira S, Kiethega GN, Burger G (2014) *Trans*-splicing and RNA editing of LSU rRNA in *Diplonema* mitochondria. Nucleic Acids Res 42(4):2660–2672. https://doi.org/10.1093/nar/gkt1152

Vidal S, Curran J, Kolakofsky D (1990) Editing of the Sendai virus P/C mRNA by G insertion occurs during mRNA synthesis via a virus-encoded activity. J Virol 64(1):239–246

Visomirski-Robic LM, Gott JM (1995) Accurate and efficient insertional RNA editing in isolated *Physarum* mitochondria. RNA 1(7):681–691

Visomirski-Robic LM, Gott JM (1997a) Insertional editing in isolated *Physarum* mitochondria is linked to RNA synthesis. RNA 3(8):821–837

Visomirski-Robic LM, Gott JM (1997b) Insertional editing of nascent mitochondrial RNAs in *Physarum*. Proc Natl Acad Sci U S A 94(9):4324–4329

Volchkov VE, Becker S, Volchkova VA, Ternovoj VA, Kotov AN, Netesov SV, Klenk HD (1995) GP mRNA of Ebola virus is edited by the Ebola virus polymerase and by T7 and vaccinia virus polymerases. Virology 214(2):421–430

Wolf PG, Rowe CA, Hasebe M (2004) High levels of RNA editing in a vascular plant chloroplast genome: analysis of transcripts from the fern *Adiantum capillus-veneris*. Gene 339:89–97. https://doi.org/10.1016/j.gene.2004.06.018

Yin YW, Steitz TA (2002) Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase. Science 298(5597):1387–1395. https://doi.org/10.1126/science.1077464

Yokobori S, Pääbo S (1995) Transfer RNA editing in land snail mitochondria. Proc Natl Acad Sci U S A 92(22):10432–10435

Yoshinaga K, Iinuma H, Masuzawa T, Uedal K (1996) Extensive RNA editing of U to C in addition to C to U substitution in the *rbcL* transcripts of hornwort chloroplasts and the origin of RNA editing in green plants. Nucleic Acids Res 24(6):1008–1014

Zauner S, Greilinger D, Laatsch T, Kowallik KV, Maier UG (2004) Substitutional editing of transcripts from genes of cyanobacterial origin in the dinoflagellate *Ceratium horridum*. FEBS Lett 577(3):535–538. https://doi.org/10.1016/j.febslet.2004.10.060

Chapter 9
Requirement of Various Protein Combinations for Each C-to-U RNA Editosome in Plant Organelles

Mizuki Takenaka, Anja Jörg, Matthias Burger, and Sascha Haag

Abstract In flowering plants, RNA editing converts several hundreds of organelle cytidines to uridines. Targeted cytidines are recognized by PLS class pentatricopeptide repeat (PPR) proteins, which bind RNA sequences upstream of the C targets in a sequence-specific manner. In the past several years, different types of proteins have been identified as RNA editing factors, including multiple organellar RNA editing factors/RNA editing factor interacting proteins (MORFs/RIPs), organelle RNA recognition motif (ORRM) proteins, organelle zinc finger (OZ) proteins, a P class PPR protein NUWA, short DYW proteins, and protoporphyrinogen oxidase 1 (PPO1). These proteins seem to contribute to individual RNA editing complexes in a different manner. Despite many key players for the assembly of editosomes having been revealed, the complete mechanism of the editing machinery including the deaminase enzymatic activity is still unclear. Plant editosomes are highly diverse not only due to the PLS class PPR proteins they contain but also in other components that are present. In this review, we introduce the recent progress in the field and discuss possible functions of each component in RNA editosomes in plant mitochondria and chloroplasts.

9.1 Introduction

RNA editing is a posttranscriptional process that alters the RNA sequence from that of DNA-encoded information. In terrestrial plants, RNA editing converts specific cytidines (Cs) to uridines (Us) at specific positions in plastid and mitochondrial transcripts (Hiesel et al. 1989; Gualberto et al. 1989; Covello and Gray 1989). While

M. Takenaka (✉)
Molekulare Botanik, Universität Ulm, Ulm, Germany

Department of Botany, Graduate School of Science, Kyoto University, Kyoto, Japan
e-mail: mizuki.takenaka@pmg.bot.kyoto-u.ac.jp

A. Jörg · M. Burger · S. Haag
Molekulare Botanik, Universität Ulm, Ulm, Germany

J. Cruz-Reyes, M. Gray (eds.), *RNA Metabolism in Mitochondria*, Nucleic Acids and Molecular Biology 34, https://doi.org/10.1007/978-3-319-78190-7_9

posttranscriptional mRNA editing is performed only as C-to-U alterations in flowering plants (Giegé and Brennicke 1999), reverse reaction (U-to-C) RNA editing also frequently occurs in ferns, mosses, and lycopods besides many C-to-U changes (Grewe et al. 2009; Kugita et al. 2003; Knie et al. 2016). The number of editing events comprises several hundred to thousands in vascular plants (Kugita et al. 2003; Bentolila et al. 2013; Oldenkott et al. 2014; Takenaka et al. 2013a), whereas the moss *Physcomitrella* has only 13 sites (Rüdinger et al. 2009). Editing often restores functionally conserved codons and creates start or stop codons that are necessary for normal expression of organellar gene functions (Chateigner-Boutin and Small 2010; Fujii and Small 2011; Knoop 2011; Shikanai 2006). Indeed, many RNA editing events have been shown to be essential for the proper function of mitochondria or plastids and optimal growth of plants (Kotera et al. 2005; Chateigner-Boutin et al. 2008; Sung et al. 2010; Zhu et al. 2012; Sosso et al. 2012; Stein et al. 1991; Zhou et al. 2008; Glass et al. 2015).

The C-to-U conversion seems to be a simple deamination reaction in plants. EDTA and EGTA severely inhibit the editing activity in mitochondrial in vitro RNA editing assays (Takenaka and Brennicke 2003), and in vitro RNA editing with chloroplast lysate is sensitive to zinc chelators, suggesting that zinc ion is an essential cofactor as in other deaminase systems (Hegeman et al. 2005). In animals, two deaminase-mediated RNA editing mechanisms have been identified, C-to-U editing by apolipoprotein B mRNA editing enzyme (APOBEC) and adenosine-to-inosine (A-to-I) editing by adenosine deaminase acting on RNA (ADAR) (Sommer et al. 1991; Teng et al. 1993; Mehta and Driscoll 1998, 2002; Bass and Weintrau 1988; Kim et al. 1994a, b; Melcher et al. 1996; Wagner et al. 1989). APOBEC- and ADAR-mediated editing systems create multiple proteins from differentially edited transcripts and regulate gene function. On the other hand, so far there is no evidence that supports a similar conditional or developmental gene regulatory function in plants (Takenaka et al. 2013a; Fujii and Small 2011; Shikanai 2015). Although unedited or partially edited transcripts may be available for plant mitochondrial ribosomes, there is no clear evidence of useful protein varieties being synthesized from incompletely edited transcripts. The majority of plant organellar proteins that accumulate derive from completely edited transcripts, likely due to instability of the aberrant proteins translated from transcripts lacking required RNA editing events (Grohmann et al. 1994; Lu and Hanson 1994; Lu et al. 1996; Phreaner et al. 1996). Therefore, the main function of RNA editing in plants seems to be a correction mechanism to compensate for T-to-C mutations inherited at the level of the organelle genome.

Editing events have been gained and lost during plant evolution. An edited C in one species is sometimes encoded as T in another (Takenaka et al. 2013a; Shikanai 2015; Bock 2000). RNA editing would not be needed if all detrimental T-to-C mutations were changed back to Ts, as observed in Marchantiaceae liverworts (Oda et al. 1992; Salone et al. 2007). However, once many mutated sites rescued by RNA editing have been fixed, converting all of them becomes increasingly unlikely, resulting in the maintenance of the RNA editing machinery (Fujii and Small 2011; Tillich et al. 2006; Gray 2012). Nevertheless, overall, RNA editing events seem to

have gradually decreased during evolution of the plant lineage; in general, lower plants and gymnosperms exhibit a larger number of target sites than angiosperms (Kugita et al. 2003; Oldenkott et al. 2014; He et al. 2016; Guo et al. 2016).

Nucleotides to be edited have to be recognized and targeted within the multitude of C nucleotides in RNA molecules. Investigations in vivo using a *trans*-plastidic approach (Bock et al. 1996, 1994; Lutz and Maliga 2007; Bock and Koop 1997), in vitro (Hegeman et al. 2005; Neuwirt et al. 2005; Verbitskiy et al. 2006, 2008; Hirose and Sugiura 2001), and *in organello* (Farré and Araya 2001; Farré et al. 2001; Blanc et al. 1995; Staudinger et al. 2005) have revealed that the *cis*-elements for C-to-U RNA editing encompass 20–25 nucleotides upstream (5′) of the editing site. However, the crucial *cis*-elements contain no commonly observed sequence motifs, implying that an individual editing site requires unique sequence recognition machinery. Indeed, it has been revealed that plant RNA editosomes include distinct RNA-binding proteins, PPR (pentatricopeptide repeat) proteins, that selectively recognize various RNA sequences located upstream of the C to be edited (Lurin et al. 2004; Cheng et al. 2016).

The discovery of a recognition code for specifying the target RNA sequences by PPR motifs, termed the "PPR code," revealed the basic mechanism for the specific selection of RNA editing sites in plant organelles (Yagi et al. 2013; Takenaka et al. 2013b; Barkan et al. 2012). All of the currently known *trans*-factors needed for RNA editing in mitochondria and chloroplasts are nucleus-encoded (Zeltz et al. 1993; Halter et al. 2004). Starting with the discovery of MORF/RIP editing factors (Takenaka et al. 2012; Bentolila et al. 2012), great progress has been made in identifying other protein components of the editosome (Shikanai 2015; Sun et al. 2016; Ichinose and Sugita 2017), now including the RRM-containing proteins (Tillich et al. 2009; Sun et al. 2013), OZ1 (Sun et al. 2015), and a P class PPR protein called NUWA (Guillaumot et al. 2017; Andrés-Colás et al. 2017; He et al. 2017). Interestingly, differences in RNA editosomes are not only provided by the PLS-type PPR recognition factors involved but also by other components. This review summarizes the recent progress in the analysis of the plant organellar editing mechanism, with a focus on the novel identified RNA editing factors required for a large number of sites that suggests an unexpected complexity of plant editosomes.

9.2 PLS-Type PPR Proteins

The first RNA editing required factor was identified for an RNA editing event in plastids in 2005 by tracing a disturbed function of the plastid NADH dehydrogenase (Kotera et al. 2005). The editing event affected creates an AUG translational start from the genomic ACG codon of the *ndhD* gene. Consequently, without this editing event, the NDHD protein is not synthesized, and the NADH dehydrogenase complex cannot be functionally assembled in the mutant. The first factor for editing events in mitochondrial mRNAs was identified by the mapping of ecotype-specific editing variants (Zehrmann et al. 2009). This protein, MEF1, is required for three editing

sites in *rps4*, *nad2*, and *nad7* transcripts in mitochondria. Up to now more than 50 proteins that belong to the same PPR gene family have been reported to be nucleus-encoded factors required for editing at one or few sites in mitochondria as well as in plastids. The PPR protein family has extraordinarily increased in the plant kingdom, with more than 400 members in a typical flowering plant (Fujii and Small 2011; Cheng et al. 2016; Lurin et al. 2004; Barkan and Small 2014; Lightowlers and Chrzanowska-Lightowlers 2013; Herbert et al. 2013). Almost all of the PPR proteins are targeted to the two genome-containing organelles, mitochondria and plastids, where these factors are involved in various RNA processing steps (Shikanai and Fujii 2013; Schmitz-Linneweber and Small 2008).

Each PPR member possesses multiple 35 amino acid repeats (PPR motifs) in tandem, which recognize specific RNA sequences in a combinatorial pattern (Fig. 9.1a) (Takenaka et al. 2013b; Barkan et al. 2012; Yin et al. 2013). The PPR proteins involved in intron removal and exon splicing, endonucleolytic processing, RNA stability, and access to translation are mostly grouped together as P class PPR proteins, which consist of tandem repeats of 35 amino acid P motifs. Those required for C-to-U RNA editing belong to the PLS class, which contains an N-terminal PPR domain with basically the repetition of a unit of canonical (P), longer (L) and shorter (S) PPR motifs, and a few additional domains at the C-terminus, the extension (E) domain, the E+ domain, and the DYW domain (Takenaka et al. 2013a; Lurin et al. 2004; Shikanai and Fujii 2013; Schmitz-Linneweber and Small 2008) (Fig. 9.1a). In *Arabidopsis thaliana*, about 190 PPR members contain E domains. Among them, approximately 80 have DYW domains in addition to EE+ domains at the C-terminus (Cheng et al. 2016; Lurin et al. 2004).

How PPR domains bind specific RNA sequences can be explained by the "PPR code" (Fig. 9.2), in which amino acid combinations at two positions in each PPR motif (6 and 1′ in the original paper but redefined as 5 and 35 after structural analysis of PPR proteins) (Yagi et al. 2013; Takenaka et al. 2013b; Barkan et al. 2012) are highly correlated with their corresponding nucleotide. PPR domains in PLS-type editing factors bind from four nucleotides upstream of the target editing site to further upstream with a one-motif to one-nucleotide correspondence manner. The current PPR code conforms to the nucleotide specificity of most P and S motifs. Certain amino acids at the fifth position provide some nucleotide preference also to L motifs, though the specificity is not as strong as that observed in P and S motifs (Yagi et al. 2013; Takenaka et al. 2013b). Combinatorial nucleotide specificity patterns of PPR motifs confer sequence-specific RNA binding to the PLS-type PPR proteins.

E domains have been shown to be essential for editing (Hayes et al. 2013; Okuda et al. 2007, 2009; Takenaka 2010). However, revealing their exact function requires further analysis. Recent redefinition of the E domain defined two tandem PPR-like structures (E1 and E2 in Fig. 9.1c), suggesting that they might have RNA binding ability similar to PPR motifs, though so far no RNA binding has been experimentally proven (Cheng et al. 2016; Okuda et al. 2014). Another possible function of the domain could be to supply an interacting interface for other cofactors in RNA editing complexes, e.g., MORF proteins. E+ domains are located between E and DYW domains (Lurin et al. 2004) and are annotated as such because a truncation of the C-terminus of PLS-type RNA editing factors often occurs preceding the E+ domain.

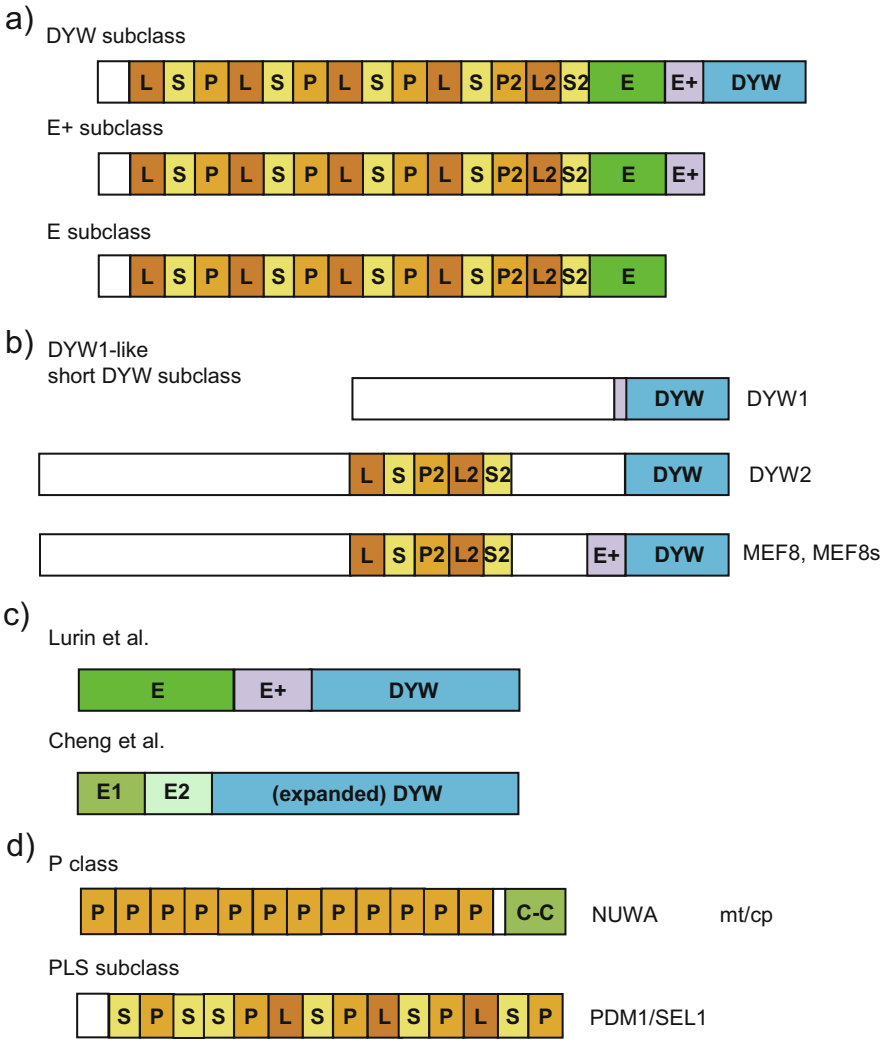

Fig. 9.1 Schematic representation of PPR (pentatricopeptide repeat) proteins involved in RNA editing. PPR proteins comprise tandem repeats of approximately 35 amino acids, forming double helix motifs. (**a**) PLS class. PPR domains in PLS class PPR proteins consist of P (canonical), L (long), and S (short) motifs. The P, L, and S motifs often form a cluster. P2, L2, and S2 are variants of P, L, and S motifs, respectively. Additional C-terminal domains have two main annotations—the extension (E) domain, E+, and the DYW domain. E and E+ subclass PPR editing factors do not have the DYW domain, and approximately 100 E or E+ subclass PPRs are encoded in *Arabidopsis*, while the DYW subclass contains about 80 proteins. (**b**) Short DYW-type PPR proteins. DYW1, DYW2, MEF8, and MEF8S have only 4–6 PPR motifs, a truncated E domain, and a DYW domain. (**c**) Definition of C-terminal domains in RNA editing factors. Original definition by Lurin et al. (2004) suggested an E, an E+, and a DYW domain, whereas a more recent definition by Cheng et al. (Cheng et al. 2016) proposed two PPR-like structures, E1 and E2, and expanded DYW domain. (**d**) NUWA and PPR596 belong to the P class, which consists of tandem repeats of canonical (35 amino acids) PPR motifs. PDM1/SEL1 has a PPR domain consisting of P, L, and S motifs but no E or DYW domains

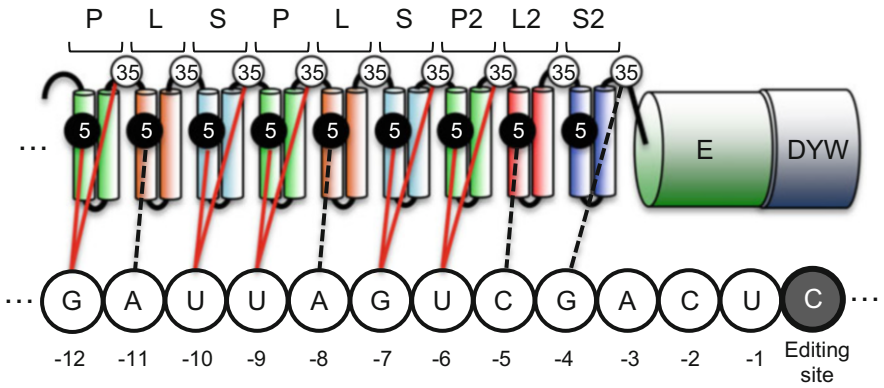

Fig. 9.2 PPR code—How PLS PPR proteins recognize the *cis*-elements for RNA editing sites. The PLS-type RNA editing PPR proteins are extended at their C-termini by E and often also by DYW domains. Different from P-type PPR proteins, the RNA editing PPR proteins contain alternating P-L-S-type elements. The P2, L2, and S2 are variants of P, L, and S motifs, respectively. The amino acid identities at positions 5 and 35 in each P, P2, and S motif are highly correlated with nucleotide identities (solid lines). Similar tendencies are also observed in some L, L2, and S2 motifs. However, their specificity is not as strong as seen in P and S motifs (broken lines)

The DYW domain shows some similarity to the cytidine deaminase domain found in other proteins (Salone et al. 2007). Two groups independently demonstrated zinc ion binding of the two conserved motifs (HXE and CXXC) in the DYW domain, which is essential for typical cytidine deaminases (Hayes et al. 2013, 2015; Boussardon et al. 2014). Therefore, this domain has been proposed to provide the C-to-U catalytic activity in the editing reaction (Salone et al. 2007). Some DYW domains have been discovered in protists (Rudinger et al. 2011; Knoop and Rüdinger 2010), and at least some of these protist species also have the plant-type C-to-U RNA editing in mitochondria, supporting the hypothesis that the DYW domain catalyzes the cytidine deaminase reaction. However, so far a deaminase activity of recombinant DYW-containing PPR editing factors has not been demonstrated in in vitro RNA editing systems, which have been established with mitochondrial or chloroplast lysates and in vitro transcribed RNA (Hegeman et al. 2005; Hirose and Sugiura 2001; Hayes and Hanson 2007; Takenaka et al. 2007; Nakamura and Sugita 2008).

9.3 Short DYW-Type PPR Proteins

Deletion of DYW domains in certain PPR proteins abolishes their editing function, suggesting that they are essential in the PPR editing factors (Zehrmann et al. 2010; Wagoner et al. 2015). However, elimination of DYW domains often does not affect the editing activity of many PPR proteins (Okuda et al. 2009, 2010; Zehrmann et al. 2010), in agreement with the fact that about half of the RNA

editing factors have no DYW domain (E or E+ subclass in Fig. 9.1a). If DYW domains contribute catalytic activity, how do E class PPR proteins work as editing factors? In 2012, DYW1, which consists of a DYW domain and a partial E domain but no PPR motifs (Fig. 9.1b), was found to be necessary for editing at a site in the chloroplast *ndhD* mRNA, which is recognized by an E subclass PPR protein, CRR4 (Fig. 9.1a) (Boussardon et al. 2012). This observation raises the possibility that the other edited Cs recognized by E or E+ subclass PPR proteins recruit DYW-containing PPR proteins to gain the deaminase activity. As the best candidates for providing DYW domains in *Arabidopsis*, five DYW domain-containing PPR proteins with several PPR motifs and a degenerate E domain have been proposed (Boussardon et al. 2012; Verbitskiy et al. 2012). Recently, DYW2, one of the members of the short-type PPR proteins, was shown to be required for more than 100 RNA editing sites (Guillaumot et al. 2017; Andrés-Colás et al. 2017). Interestingly, many of the DYW2-affected editing sites are also known as targets of E+ subclass PPR proteins, suggesting that DYW2 is required for providing DYW domains to all E+ subclass PPR proteins. Truncated C-terminal domains in DYW1 or DYW2 seem to complement the missing C-terminal domains of the long PPR partners (Fig. 9.1a, b). So far the partners for E subclass PPRs have not been identified, but truncated E domains in some short DYW proteins perfectly compensate for the missing C-terminus of E subclass PPR proteins, implying a similar scenario for E subclass PPRs. Another short DYW member, MEF8, shows reduced RNA editing at several sites in the knockout line (Verbitskiy et al. 2012; Diaz et al. 2017). Interestingly, the double knockout of MEF8 and its close homolog MEF8S show embryo lethality, suggesting redundancy of their function (Verbitskiy et al. 2012). As well, the MEF8 and MEF8S proteins may provide the enzymatic activity to an editosome carrying a site-specific E subclass PPR, though so far no E subclass PPR proteins that edit affected sites in *mef8* mutants have been reported (Verbitskiy et al. 2012; Diaz et al. 2017). The cooperation between E or E+ subclass PPR proteins and short-type PPR proteins indicates that an RNA editing complex likely includes at least one DYW domain-containing protein. This and other already mentioned data strongly suggest that the DYW domain is sufficient as a catalytic domain in the plant RNA editosome (Hayes et al. 2013, 2015; Boussardon et al. 2014). However, it is hard to deny the possibility that another, still missing, protein is required for performing the C-to-U deamination.

Until recently, all identified editing factors belonged to the PLS class PPR family (Kotera et al. 2005; Shikanai 2015; Takenaka 2014). Accordingly, a simple RNA editosome model comprising two types of PPR proteins was proposed: an E or E+ subclass PPR protein recognizes the RNA sequence and another DYW domain-containing protein supplies the C-to-U deaminase activity. However, discovery of other essential components in the editosome, the MORF/RIP family, the ORRM family, the OZ family, and the P class PPR protein family indicates a more complex model of RNA editosomes in plant organelles (Table 9.1) (Takenaka et al. 2012; Bentolila et al. 2012; Sun et al. 2013, 2015; Guillaumot et al. 2017; Andrés-Colás et al. 2017; Shi et al. 2015, 2016).

Table 9.1 RNA editing factors in plant organelles other than typical E or DYW class site-specific factors

Protein family	Protein	Alias	Accession # in *Arabidposis*	Subcellular localization
Short DYW	DYW1		AT1G47580	Chloroplasts
Short DYW	DYW2		AT2G15690	Mitochondria/ chloroplasts
Short DYW	MEF8		AT2G25580	Mitochondria
Short DYW	MEF8S		AT4G32450	Mitochondria
P class PPR	NUWA	EMB1796	AT3G49240	Mitochondria/ chloroplasts
P class PPR	PPR596		AT1G80270	Mitochondria
PLS subclass PPR	PDM/ SEL1		AT4G18520	Chloroplasts
MORF/RIP	MORF1	RIP8	AT4G20020	Mitochondria
MORF/RIP	MORF2	RIP2, DAL	AT2G33430	Chloroplasts
MORF/RIP	MORF3	RIP3	AT3G06790	Mitochondria
MORF/RIP	MORF8	RIP1	AT3G15000	Mitochondria/ chloroplast
MORF/RIP	MORF9	RIP9	AT1G11430	Chloroplasts
RRM	CP31A	RBP31	AT4G24770	Chloroplasts
RRM	CP31B		AT5G50250	Chloroplasts
RRM	ORRM1		AT3G20930	Chloroplasts
RRM	ORRM2		AT5G54580	Mitochondria
RRM	ORRM3	GR-RBP3, RBGA7	AT5G61030	Mitochondria
RRM	ORRM4	GR-RBP5, RBGA2	AT1G74230	Mitochondria
RRM	ORRM5	GR-RBP2, GRP2, RBGA5	AT4G13850	Mitochondria
RRM	ORRM6		AT1G73530	Chloroplasts
OZ	OZ1	VAR3	AT5G17790	Chloroplasts
	PPO1	PPOX1	AT4G01690	Chloroplasts
	OCP3		AT5G11270	Chloroplasts
	HEMC	RUG1, RUGOSA1	AT5G08280	Chloroplasts

9.4 MORF Proteins

Multiple organellar RNA editing factor 1 (MORF1) has been identified through genetic screening for mitochondrial editing defects in EMS mutant populations of *Arabidopsis* (Takenaka et al. 2012). The *morf1-1* mutant, in which a single amino acid substitution occurs in a conserved sequence, shows reduced RNA editing extent at more than 50 sites in mitochondria. Another member of the MORF proteins, RIP1 (= MORF8), was isolated as one of the cofactors of a PPR RNA editing factor, RARE1 (Bentolila et al. 2012). In a *rip1 Arabidopsis* mutant, the editing efficiency at over 400 sites in mitochondria and 11 sites in chloroplasts was reduced (Bentolila et al. 2012, 2013). All MORF proteins in *A. thaliana* are imported into mitochondria

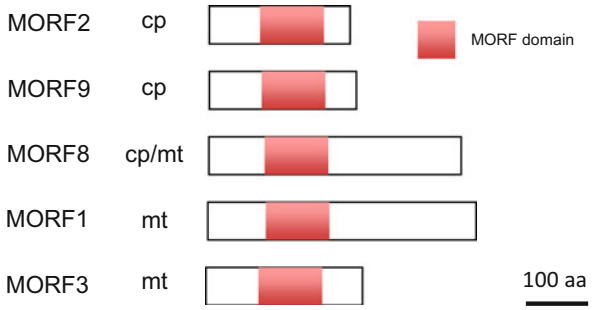

Fig. 9.3 Of the nine MORF proteins in *Arabidopsis*, five MORF (multiple organellar RNA editing factor) proteins play a major role in plant RNA editing. All MORF proteins have a central conserved motif (MORF domain), while MORF1 and MORF8 have an extended C-terminus with unknown function

and/or chloroplasts and contain a highly conserved central region of approximately 100 amino acids termed the MORF domain (Fig. 9.3). In contrast to genes encoding PLS-type PPR proteins, mutagenesis of a single MORF/RIP gene affects dozens or hundreds of editing sites (Bentolila et al. 2013). Among the nine members of the MORF/RIP family in *Arabidopsis*, five show a significant effect on mitochondria and/or plastid editing while the rest of the family members have minor or no effects on editing (Bentolila et al. 2013; Takenaka et al. 2012) (Table 9.1). In plastids, MORF2 and MORF9 are required for almost all editing sites (Bentolila et al. 2013; Takenaka et al. 2012), and the knockout of each gene leads to impaired chloroplast development (Takenaka et al. 2012; Chatterjee et al. 1996; Bisanz et al. 2003). In mitochondria, MORF8/RIP1 affects over 400 sites and MORF1 and MORF3 around 50–60 sites.

Besides the conserved MORF domain, MORFs do not have any annotated domain, but MORF1, MORF4, and MORF8 contain an extended C-terminus (Fig. 9.3) (Bentolila et al. 2013; Takenaka et al. 2012). Recent structural analysis of the MORF domain suggested its structural similarity to an N-terminal ferredoxin-like domain (NFLD), which confers RNA substrate positioning in bacterial 4-thiouracil tRNA synthetases, implying direct RNA contacts of MORF proteins during RNA editing, though this has to be experimentally proven (Haag et al. 2017). MORFs interact with the PLS-type PPR editing factors via a connection to the PPR motifs or other regions (Takenaka et al. 2012; Bentolila et al. 2012; Bayer-Császár et al. 2017). Some editing events can be affected by two or more different MORF proteins, suggesting that different combinations of MORF editing factors are required for distinct editing events (Bentolila et al. 2013; Takenaka et al. 2012). Indeed, such complex interaction patterns were observed among MORF members through protein-protein interaction studies (Takenaka et al. 2012; Zehrmann et al. 2015) (Table 9.2). Furthermore, structural studies of MORF1 and MORF9 proteins confirmed the homodimerization of MORF domains. Interestingly, MORF1 shows a possible tetramer formation, suggesting more complex regulation of RNA editing through homo- and hetero-interactions between several MORF proteins (Haag et al. 2017). As

Table 9.2 Analyzed protein–protein interaction between known RNA editing factors

Chloroplasts													
	E or DYW class PPR	DYW1	DYW2	NUWA	MORF2	MORF8	MORF9	ORRM1	ORRM6	OZ1	PPO1	HEMC	PDM/SEL
E or DYW class PPR	Na	Y[c]	Y[bc]	Y[bc]	Y[abc]	Y[abc]	Y[abd]	Y[a]	No evidence	Y[ac]	No evidence	Y[ac]	Y[a]
DYW1	–	Na	Na	Na	Na	Na	Na	Na	Na	Na	Na	Na	Na
DYW2	–	–	N[b]	Y[abc]	Na	Na	Na	Na	Na	Na	Na	Na	Na
NUWA	–	–	–	Y[b]	Y[b]	Na	Na	Na	Na	Na	Na	Na	Na
MORF2	–	–	–	–	Y[abd]	Y[abcd]	Y[abd]	Y[a]	Y[a]	Y[a]	Y[ac]	N[a]	Y[a]
MORF8	–	–	–	–	–	Y[abd]	Y[ad]	Y[a]	Y[a]	Y[a]	Y[a]	Y[ad]	Y[a]
MORF9	–	–	–	–	–	–	Y[abd]	N[a]	Y[a]	N[a]	Y[ac]	N[a]	Y[a]
ORRM1	–	–	–	–	–	–	–	Na	N[a]	Y[a]	Na	N[a]	Na
ORRM6	–	–	–	–	–	–	–	–	Y[a]	Y[a]	Na	Na	Na
OZ1	–	–	–	–	–	–	–	–	–	Y[a]	Na	Na	Na
PPO1	–	–	–	–	–	–	–	–	–	–	Na	N[a]	Na
HEMC	–	–	–	–	–	–	–	–	–	–	–	Na	Na
PDM/SEL1	–	–	–	–	–	–	–	–	–	–	–	–	Na

Mitochondria												
	E or DYW class PPR	DYW2	MEF8	MEF8S	NUWA	MORF1	MORF3	MORF8	ORRM2	ORRM3	ORRM4	ORRM5
E or DYW class PPR	Na	Y[a]	Na	Na	Y[a]	Y[abcd]	Y[abd]	Y[ab]	No evidence	No evidence	No evidence	No evidence
DYW2	–	Y[b]	Na	Na	Y[abc]	Y[c]	Na	Na	Na	Na	Na	Na
MEF8	–	–	Na	Na	Na	Y[a]	Y[a]	Y[a]	Na	Na	Na	Na
MEF8S	–	–	–	Na	Na	Y[a]	Y[a]	Y[a]	Na	Na	Na	Na
NUWA	–	–	–	–	Y[b]	Y[c]	Na	Na	Na	Na	Na	Na
MORF1	–	–	–	–	–	Y[abd]	Y[abd]	Y[abcd]	Na	Na	Na	Na
MORF3	–	–	–	–	–	–	Y[dbd]	Y[abd]	N[a]	N[a]	Na	Na
MORF8	–	–	–	–	–	–	–	Y[ad]	N[a]	N[a]	Y[a]	N[a]
ORRM2	–	–	–	–	–	–	–	–	N[a]	Y[a]	N[a]	Y[a]
ORRM3	–	–	–	–	–	–	–	–	–	Y[a]	Y[a]	Y[a]
ORRM4	–	–	–	–	–	–	–	–	–	–	Y[a]	Y[a]
ORRM5	–	–	–	–	–	–	–	–	–	–	–	N[a]

[a]Yeast two-hybrid assay
[b]In vivo biomolecular fluorescence complementation assay
[c]In vivo pull-down assay
[d]In vitro pull-down assay
[e]Interaction results summarized here are based on Sun et al. (2015), Guillaumot et al. (2017), Andrés-Colás et al. (2017), Boussardon et al. (2012), Shi et al. (2015, 2016, 2017), Bayer-Császár et al. (2017), Zehrmann et al. (2015), Hackett et al. (2017), Zhang et al. (2014, 2015), Huang et al. (2017)
[f]Individual factors are listed except for the E- and DYW-type PPR proteins. *Y* interaction, *N* no interaction, *Na* not analyzed, no evidence, not possible to conclude from the results with only a few E- and DYW-type PPR proteins

components of RNA editosomes, specific homomers of MORFs are required for some sites, but a specific heterogeneous combination can be more active for others. In addition to the MORF-MORF interactions, MORF proteins preferably interact with PLS class PPR proteins including the short DYW class, which is likely to be required for the function of MORF proteins (Bayer-Császár et al. 2017). A requirement for a specific MORF combination in a PPR-MORF interaction is shown for the MEF13 protein, whose target editing sites are also negatively affected in *morf3* and *morf8* mutant lines (Glass et al. 2015). Interaction between the MEF13 and MORF3 proteins is enhanced only in the presence of MORF8 but not MORF1, in a yeast three-hybrid system.

It has been hypothesized that MORF dimers serve as a bridge between the PLS-type PPR editing factor and the deaminase activity in the editosome (Fig. 9.4a) (Takenaka 2014; Zehrmann et al. 2015), since two or more MORF proteins affect the same editing sites and MORF proteins form different homo- and heterodimers (Table 9.2). Yeast two-hybrid and pull-down assays between MORF1 and MEF21, both of which are required for editing at position 257 in *cox3* mRNA (*cox3*-257), suggest that the MORF1 protein strongly binds to the E domain of MEF21 (Bayer-Császár et al. 2017). Therefore the E domain may be a main hub for PPR-MORF interaction.

Another structural analysis for the MORF9 protein with an artificially designed PLS repeat protein indicates that MORF9 directly associates with PPR motifs (Yan et al. 2017). In a REMSA assay, the MORF9 protein enhances the interaction between the artificial PPR protein and its target RNA. MORF9 associates with P, L, or S motifs, but only converts an angle between the two helixes in an L motif to angles similarly observed in P- and S-type motifs. Such structural changes by the MORF9 protein may enable L motifs to contribute to sequence-specific RNA binding, which is generally observed in P and S motifs (Fig. 9.4b) (Yan et al. 2017). Although the enhancement of PPR-RNA affinity is so far the most reasonable explanation for a role of MORF proteins, it is unclear how specific MORF-MORF interactions contribute to such a scenario. The absence of E and DYW domains in the artificial PPR protein employed for the structural analysis of the MORF9-PPR complex leaves room for speculations about other MORF functions in RNA editosomes (Yan et al. 2017). Furthermore, the presence of other MORF-interacting cofactors that will be mentioned in the following sections renders it more complicated to elucidate how MORF proteins take part in plant RNA editosomes.

While the MORF family is widespread among angiosperms, genes encoding MORFs are missing in some ferns and moss species that also have organelle RNA editing (Takenaka et al. 2012, 2013a; Bentolila et al. 2012; Luo et al. 2017). So far there is no clear explanation why angiosperm RNA editing requires these proteins while editing in other plants does not. The PPR domains in the moss and fern editing factors are generally longer and contain more S-motif repeats than angiosperm ones do, suggesting that these PPR domains inherently possess higher affinities for their target RNA sequences (Cheng et al. 2016). Therefore, an enhancing effect for RNA affinity by MORF-L motif interaction may not be necessary. The absence of MORF proteins in the moss *Physcomitrella* may be explainable by the presence of DYW

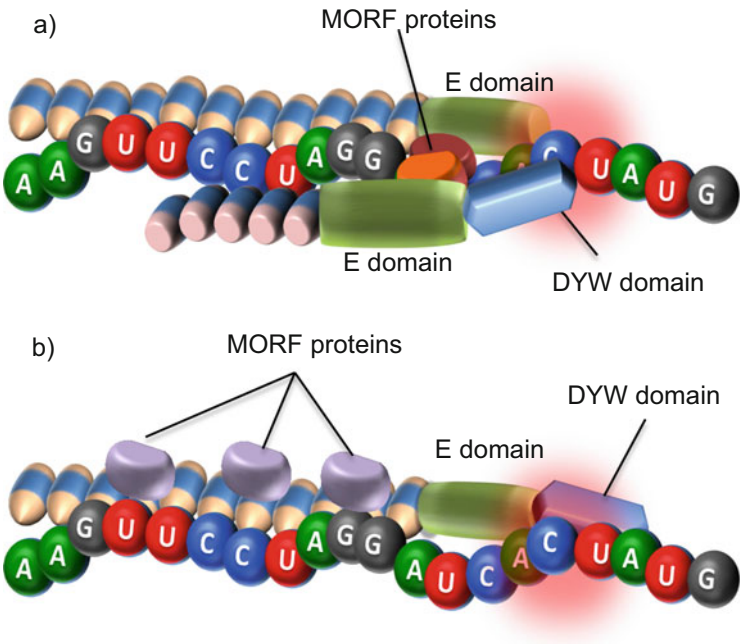

Fig. 9.4 Proposed roles of MORF proteins in RNA editosomes. (**a**) MORF proteins form a homo- and heterodimer, which supports interaction between E or E+ PPR proteins and a short DYW protein. These models may occur simultaneously. (**b**) MORF proteins bind to PPR motifs. The MORF association changes the structure of the L motif and consequently enhances affinities to target RNA sequence

domains in all PLS-type PPR factors in this plant (Rüdinger et al. 2009; Sugita et al. 2013). If the MORF proteins are required for supporting a complex formation with DYW subclass and E subclass PPRs, they would not be necessary in *Physcomitrella*. On the other hand, this scenario does not fit for the spikemoss *Selaginella*, which has many E class PPR proteins (Fujii and Small 2011). In this plant, other as yet unknown proteins may substitute for the bridge function of MORF proteins.

9.5 RRM Motif-Containing Proteins

CP31 proteins bind several RNA substrates for RNA editing in chloroplasts, and immunodepletion of them from a chloroplast lysate decreases in vitro editing activity at some sites (Hirose and Sugiura 2001; Miyamoto et al. 2002). CP31A and B, two closely related proteins, have two RRM (RNA recognition motif) domains (Fig. 9.5a) and are categorized as chloroplast ribonucleoproteins (cpRNPs). Knock-out mutant lines of CP31A show lower editing extents at multiple plastid sites. On the other hand, a *cp31b* mutant shows no effect on editing. However, a double

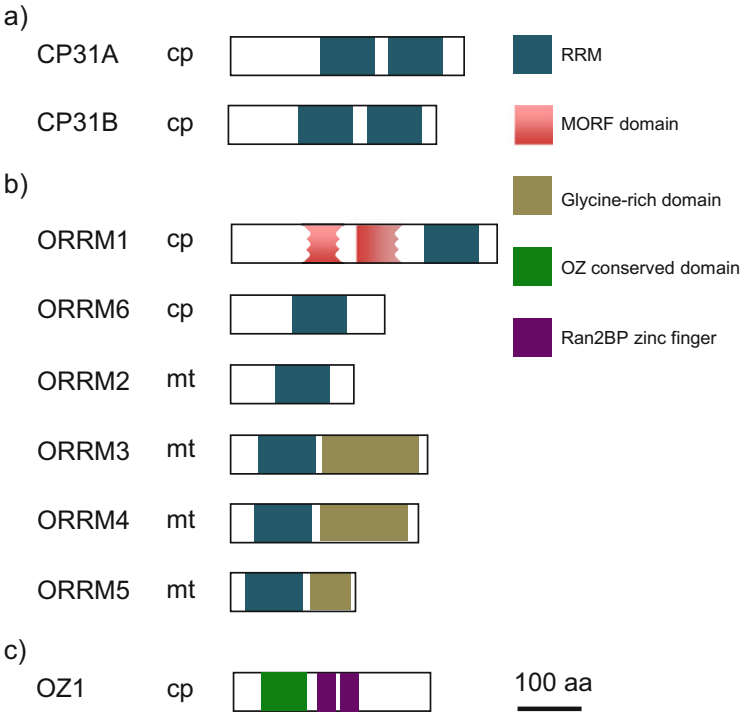

Fig. 9.5 RRM-containing proteins and OZ1 are required for RNA editing in plant organelles. (**a**) Six organelle RNA recognition motif (ORRM) proteins have been shown to be involved in RNA editing. ORRM1 has two degenerate MORF/RIP domains and one RRM motif, while other ORRMs do not have MORF domains. ORRM3 and ORRM5 have glycine-rich regions. (**b**) CP31A and CP31B have two RRM motifs. (**c**) There are four members in the organelle zinc finger (OZ) protein family. So far, OZ1 is the only protein that has been identified as an RNA editing factor. OZ1 has a region conserved within other family members in addition to two RanBP2-type zinc fingers

knockout of CP31A and CP31B reduces the editing extent at three additional sites in addition to the affected sites in *cp31a* mutant (Tillich et al. 2009). Transcript abundance of chloroplast-encoded genes is greatly reduced in *cp31* mutants, suggesting that CP31 may primarily be required for controlling RNA stability (Tillich et al. 2009).

ORRM1 was initially focused on as a potential RNA editing factor due to its two truncated MORF domains in addition to an RRM domain (Sun et al. 2013) (Fig. 9.5b; Table 9.1). The loss of ORRM1 protein influences over 60% of plastid editing sites in *Arabidopsis*. Interestingly, the RRM domain of ORRM1 is solely sufficient for restoring RNA editing in an *orrm1* mutant (Sun et al. 2013). Knockout lines of the maize ortholog of ORRM1 also show reduced editing extents at multiple sites in plastids, implying a conserved editing function across species. Although the RRM is a common motif in eukaryotes, the motif in ORRM1 belongs to a distinct

clade of around 20 members in *Arabidopsis*. None of the ORRM members other than ORRM1 contains any sign of a MORF domain; instead, many contain glycine-rich regions (Sun et al. 2013) (Fig. 9.5b).

So far six ORRM proteins have been reported to be involved in RNA editing (Table 9.1) (Sun et al. 2013; Shi et al. 2015, 2016, 2017; Hackett et al. 2017). ORRM1 and ORRM6 are editing factors in plastids, whereas ORRM2, ORRM3, ORRM4, and ORRM5 are required for mitochondrial editing. Similar to MORF/RIP factors, ORRMs affect many editing sites in a site-specific manner, though ORRM4 and ORRM5 also broadly influence editing events in the same transcripts (Sun et al. 2013; Shi et al. 2015, 2016, 2017; Hackett et al. 2017). In contrast to what is observed in the case of other ORRM editing factors, *orrm5* mutant plants show increased editing extent in approximately 15% of the mitochondrial sites (Shi et al. 2017). The *orrm5* mutant also exhibits a reduced splicing efficiency of the first *nad5* intron, slower growth, and delayed flowering time. ORRMs can interact with MORF proteins and form homo- or heterodimers (Sun et al. 2013; Shi et al. 2015, 2016, 2017; Hackett et al. 2017). ORRM1 associates with PLS-type PPR editing factors, while the interactions between ORRM2–4, ORRM6, and PPR factors are likely to be mediated via MORFs or other components of the editosome (Table 9.2). ORRM5 does not interact with MORF8/RIP1 but with ORRM2–4 proteins. ORRM1 and ORRM6 can bind RNA containing *cis* elements for some editing sites (Sun et al. 2013; Hackett et al. 2017), indicating that the RRM domains in the ORRM proteins have the ability to bind RNA as well as having affinity for other proteins.

It is not clear how RRM proteins contribute to RNA editing in plant organelles. One possibility is that in binding to transcripts containing RNA editing sites, RRM proteins are required for altering the RNA secondary structure to improve the accessibility of other factors or to support the specific *cis*-element recognition by PLS PPR proteins (Tillich et al. 2009). Considering their affinities to other editing factors, RNA binding of RRM domains may not be important for the editing function of ORRM proteins. The discovery of the ORRM family proteins suggests that other RRM proteins may also be involved in plant organellar RNA editing.

9.6 OZ Proteins

OZ1 was isolated as one of the ORRM1-co-purified proteins (Sun et al. 2015). Loss of OZ1 in *Arabidopsis* leads to editing defects at 30 plastid sites, which results in a virescent (greenish) phenotype (Table 9.1) (Sun et al. 2015; Naested et al. 2004). OZ1 selectively associates with the PLS PPR specificity factors and also strongly interacts with ORRM1 but does not show direct interaction with MORFs. In addition, OZ1 is able to form a homodimer (Table 9.2) (Sun et al. 2015). There are four OZ1-like genes in *Arabidopsis*; all of which are organelle-targeted. All OZ family proteins have various numbers of RanBP2-type zinc finger domains (Fig. 9.5c). In addition, there is another conserved domain of unknown function in the four OZ proteins. The elucidation of the function of the zinc finger and the

second domain of the OZ proteins in the editosome requires further investigation. As mentioned in the above sections, a requirement for zinc ions in plant C-to-U editing has been shown. Moreover, it has also been demonstrated that the DYW domain in the PLS-type PPR editing factors, which is the best candidate for the cytidine deaminase, binds to zinc ions (Hayes et al. 2013, 2015; Boussardon et al. 2014). Although it has not yet been analyzed whether the zinc finger domains in OZ proteins actually bind Zn^{2+}, OZs may be required for supplying zinc ions to editing deaminases. The OZ family is present in many plant lineages including *Selaginella* and mosses but not in *Chlamydomonas* or *Volvox*, which have no RNA editing, suggesting a conserved function in the plant editosome (Sun et al. 2015).

9.7 P Class PPR Proteins

P class PPR proteins consist of tandem repeats of 35 amino acid P-type PPR motifs, though some of them contain additional domains (Fig. 9.1d). In contrast to PLS class PPR proteins, most P class PPR proteins have been shown to be involved in various RNA processing steps other than RNA editing (Fujii and Small 2011; Lurin et al. 2004; Barkan and Small 2014; Shikanai and Fujii 2013; Delannoy et al. 2007). PPR596 is the first reported P class PPR protein and influences RNA editing at one of the *rps3* sites in mitochondria (Fig. 9.1d). Loss of the *PPR596* gene increases editing at the *rps3*-1344 site, suggesting that this P class PPR protein negatively affects editing efficiency, probably through altering the structure of the *rps3* transcript or competing with a PLS-type PPR specificity factor for the site (Doniwa et al. 2010).

Recently, another P class PPR protein, NUWA, was isolated as a protein co-immunoprecipitated with E+ subclass PPR RNA editing factors in chloroplasts or mitochondria (Fig. 9.1d) (Guillaumot et al. 2017; Andrés-Colás et al. 2017). NUWA is involved in numerous sites in both plant organelles, and its null mutant shows an embryo-lethal phenotype. Embryo-specific expression of the NUWA gene restores embryo development, and a lower extent of editing at many mitochondrial and chloroplast sites is observed in transcripts from mature leaves (Guillaumot et al. 2017). The affected editing sites in the NUWA mutant lines largely overlap with the sites targeted by E+ subclass PPR editing factors, similarly as in mutant lines of the short DYW protein, DYW2, which is also co-immunoprecipitated with E+ subclass PPR proteins (Guillaumot et al. 2017; Andrés-Colás et al. 2017). Bimolecular fluorescence complementation analyses suggested that NUWA supports the interaction between SLO2, a mitochondrial E+ subclass PPR, and DYW2 in mitochondria (Andrés-Colás et al. 2017). The NUWA proteins are co-immunoprecipitated with MORF1 or MORF2 proteins, suggesting that this protein also associates with MORF proteins in vivo (Bayer-Császár et al. 2017). Although the NUWA protein contains 12 PPR motifs (Fig. 9.1d), it is unclear whether they contribute to specific RNA binding for target selection. Discovery of the NUWA protein implies that other P class PPR proteins may also play an important role in RNA editing in plant organelles.

9.8 Other Factors

Four additional proteins—PDM/SEL1, PPO1, OCP3, and HEMC—have been found to influence RNA editing efficiency in plastids. PDM1/SEL1 is involved in the RNA editing of the *accD* transcript in chloroplasts (Pyo et al. 2013) and the splicing of group II introns in transcripts of *trnK* and *ndhA* (Zhang et al. 2015). This protein contains a PLS-type PPR domain but no C-terminal E and DYW domains. Therefore, it is unlikely to be a typical specificity factor like E- or DYW-class PPR proteins. It may affect RNA editing through interactions with MORF2, MORF8, and MORF9 (Table 9.2) (Zhang et al. 2015).

A key enzyme for tetrapyrrole metabolism, PPO1, was shown to have a role in chloroplast RNA editing (Zhang et al. 2014). Eighteen plastid sites showed reduced editing in PPO1 mutant lines. PPO1 directly associates with plastid MORFs/RIPs but not with PLS-type PPR factors (Table 9.2). The respective editing defects can be rescued by a truncated PPO1 protein, which lacks oxidase enzymatic activity (Table 9.1) but maintains MORF binding (Zhang et al. 2014). Overexpression of cationic peroxidase 3 (OCP3) also affects editing of multiple sites in the plastid *ndhB* transcript (Table 9.1) (García-Andrade et al. 2013). Although *ocp3* mutants exhibit only a moderate reduction in editing extent at the *ndhB* mRNA, NDH activity was decreased with a concomitant increase in resistance for fungal infection (García-Andrade et al. 2013). HEMC, a porphobilinogen deaminase that operates upstream of the chlorophyll biosynthetic pathway, has been isolated as an interacting protein with AtECB2, a PLS-type PPR RNA editing factor in chloroplasts (Huang et al. 2017). Mutant lines for the HEMC gene show lower editing extents at four chloroplast RNA editing sites. This protein interacts with MORF8/RIP1 but not with other chloroplast MORF proteins, PPO1 or ORRM1 (Table 9.2) (Huang et al. 2017). Unexpected influence of these essential key metabolic enzymes on chloroplast RNA editing may suggest possible connections between the respective essential metabolic pathways and RNA editing. However, it is more likely that these bindings could simply be fortuitous, and the recruitment of these proteins to act in RNA editing is just another example of constructive neutral evolution (CNE), in which neutral protein interactions gain functional dependencies (Gray 2012; Stoltzfus 2012).

9.9 Summary: Various Forms of Plant RNA Editosomes

In plant organellar RNA editing, specific selection of particular cytidines is crucial to prevent detrimental changes in transcripts. The specificity of the recognition of upstream *cis*-elements is basically explained by the PPR code, whereby each PPR motif recognizes one or two particular nucleotides (Yagi et al. 2013; Takenaka et al. 2013b; Barkan et al. 2012; Yin et al. 2013; Gully et al. 2015; Shen et al. 2016; Kindgren et al. 2015). However, not all PPR motifs show the strict nucleotide specificities (Yagi et al. 2013; Takenaka et al. 2013b). Additionally, each PPR

motif unequally contributes to the RNA binding capacity of the entire protein (Okuda et al. 2014; Kindgren et al. 2015). Therefore, several different C targets with similar but not identical *cis*-elements can be recognized by a single PPR protein (Chateigner-Boutin et al. 2008; Glass et al. 2015; Kim et al. 2009; Hammani et al. 2009). This flexibility of PPR editing factors in site recognition might be evolutionarily important for preadaptation in advance fixation of T-to-C mutations in the organellar genome. A plant in which new T-to-C mutations can be restored via C-to-U RNA editing by pre-existing PLS PPR editing factors, even if only partially, should have an advantage. Evolutionary selection for PLS PPR editing factors should favor improved editing efficiency at a new C without loss of specificity at the original target sites.

An early hypothesis for the RNA editosome proposed the presence of one or more PPR proteins that supply an editing activity including a deaminase at the appropriate location, such as a combination of CRR4 and DYW1 or an E+ PPR protein and DYW2 (Boussardon et al. 2012). The discovery of many editing factors other than PLS-type PPR proteins indicates that there are additional varieties of editosomes in plant organelles. Most of these newly found non-PPR and P class PPR factors affect many more editing targets than any individual PLS class PPR protein, indicating that these factors are involved in a larger number of editosomes.

The functional significance of most interactions between particular non-PPR editing factors and specific PLS-type PPR proteins or between non-PPR editing factors is still unclear, though yeast two-hybrid, bimolecular fluorescence complementation, and pull-down assays indicate that some selective interactions seem to occur also in plant mitochondria and plastids (Table 9.2). Although loss of expression of *orrm1*, *oz1*, and *morf/rip* genes can sometimes induce complete absence of editing at certain sites, some members of the respective gene families seem to be functionally redundant. Removing one factor often results in moderate reduction of editing efficiency at particular C targets, which may suggest that some family members can partially compensate for editing defects in the absence of another member. For example, at the mitochondrial *cox3*-314 site, at least MEF13, DYW2, NUWA, MORF3, and MORF8 are required for efficient RNA editing. Editing of *cox3*-314 was reduced to 40% when *morf8/rip1* was mutated, while 70% editing remained in the *morf3* knockout mutant (Takenaka et al. 2012; Bentolila et al. 2012). Possibly the editosomes for *cox3*-314 are more efficient when they include a heteromer of MORF3 and MORF8, but one without MORF3 or MORF8 is still partially functional. The expression of MORF8 enhances the interaction between MEF13 and MORF3 in a yeast three-hybrid system (Glass et al. 2015), strongly supporting this hypothesis. Another example comes from a study of two ORRMs, ORRM2 and ORRM3, that share many target sites in mitochondria. While an *orrm3* mutant still shows a residual extent of editing, silencing of ORRM2 in the *orrm3* mutant further reduced the editing efficiency (Shi et al. 2015). ORRM2 and ORRM3 form heterodimers in yeast two-hybrid analyses, and ORRM3 can form a homodimer (Shi et al. 2015). An editosome-containing ORRM3 or ORRM2-ORRM3 heteromers may be substituted by an ORRM2, if ORRM3 is missing. Possibly, however, an editosome carrying only ORRM2 is not efficiently active or

not sufficiently stable. To understand the function of respective RNA editing factors in RNA editing complexes, structural analyses of each editing factor can be a direct and effective approach, as demonstrated for PPR and MORF proteins (Yin et al. 2013; Haag et al. 2017; Yan et al. 2017).

The size of minimum editosomes derived from size-exclusion chromatography experiments is stated to be about 200 kDa (Bentolila et al. 2012), which fits with the combined size of factors identified so far for a particular site, implying that possibly all of the proteins present in such editosomes have been identified. It is still possible, however, that there are other novel editosome components crucial for the editing reaction.

Despite the fact that many non-PPR editing factors affect a large number of editing sites, some sites are not affected in any *morf*, *orrm*, *nuwa*, or *oz.* mutants or in tissue-silenced lines. These sites may maintain a more simple system, e.g., with only a PLS-type PPR protein and a protein carrying the deaminase activity. Therefore, any additional auxiliary factors may not be necessary. Alternatively, these sites could be controlled by many different editing factors, so that redundancy rescues mutant phenotypes, or they require some proteins that have not yet been identified.

MORF proteins and direct homologs of ORRM proteins are not encoded in *Selaginella*, ferns, and mosses (Sun et al. 2013). All PLS-type PPR editing factors in the moss *Physcomitrella* contain DYW domains, suggesting that they do not need to recruit DYW domains in *trans* and other auxiliary factors to assemble complex editosomes (Rüdinger et al. 2009; Sugita et al. 2013). However, many PLS-type PPR proteins in *Selaginella* do not have DYW domains, indicating that both E and DYW subclass PPR proteins may be available as RNA editing factors in this plant, as in angiosperms (Fujii and Small 2011). In nonflowering vascular plants like *Selaginella*, other proteins may play a similar role to support complex formation with E or E+ subclass and DYW subclass PPR proteins. To completely understand RNA editosomes in plant organelles, reconstitution of the active editing complex using isolated recombinant proteins is indispensable.

Many editing sites are not shared among species or lineages, indicating frequent gain and loss of editing sites during plant evolution (Takenaka et al. 2013a; Shikanai and Fujii 2013). Several possible scenarios to explain gain of novel potential targets for C-to-U RNA editing system can be considered. One is simply an increase in the number of PLS-type PPR proteins that bind different RNA sequences. Acquisition of cofactors that enhance or alter RNA affinity in existing PLS-type PPR proteins should also lead to an increase in the number of potential target sites. Furthermore, separation of E and DYW domains and reassembly with different combinations would likewise be expected to result in an increase in the number of accessible sites, if each DYW domain also has a distinct sequence specificity. Recruitment of proteins supporting interaction between E and DYW subclass PPR proteins might also lead to an increase in possible target site number.

How have plants evolved such a complex RNA editing system? The concept of constructive neutral evolution (CNE) is so far the most reasonable explanation for the massively increased number of RNA editing sites in plant organelles (Gray 2012; Stoltzfus 2012). Prior existence of editing activities and a further increase in

potential target sites passively allow fixation of T-to-C mutations in the genome, especially when such T-to-C mutations occur frequently as a result of a drive toward more GC-rich genomes. The diversity of complex plant editosomes that have emerged during plant evolution is also well explained by CNE, according to which RNA editosomes are assembled from existing proteins, which already have a metabolic role in the cell. Complexity of the editosomes might evolve through a process of "presuppression," in which each editosome component gains mutations that make it required for function on other pre-existing components or processes (Gray 2012). The diversity of C-to-U RNA editosomes in plant organelles may reflect a result of independent CNE processes in each editosome.

Acknowledgments Recent work on RNA editing was supported by grants from the Deutsche Forschungsgemeinschaft [TA642/10-1, TA 642/6-1, TA 642/3-1 to MT]. Authors thank Axel Brennicke for helpful suggestions.

References

Andrés-Colás N, Zhu Q, Takenaka M, De Rybel B, Weijers D, Van Der Straeten D (2017) Multiple PPR protein interactions are involved in the RNA editing system in *Arabidopsis* mitochondria and plastids. Proc Natl Acad Sci 114:8883–8888. https://doi.org/10.1073/pnas.1705815114

Barkan A, Small I (2014) Pentatricopeptide repeat proteins in plants. Annu Rev Plant Biol 65:415–442

Barkan A, Rojas M, Fujii S, Yap A, Chong YS, Bond CS, Small I (2012) A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. PLoS Genet 8:e1002910. https://doi.org/10.1371/journal.pgen.1002910

Bass BL, Weintrau H (1988) An unwinding activity that covalently modifies its double-stranded RNA substrate. Cell 55:1089–1098. https://doi.org/10.1016/0092-8674(88)90253-X

Bayer-Császár E, Haag S, Jörg A, Glass F, Härtel B, Obata T, Meyer EH, Brennicke A, Takenaka M (2017) The conserved domain in MORF proteins has distinct affinities to the PPR and E elements in PPR RNA editing factors. Biochim Biophys Acta 1860:813–828. https://doi.org/10.1016/j.bbagrm.2017.05.004

Bentolila S, Heller WP, Sun T, Babina a M, Friso G, van Wijk KJ, Hanson MR (2012) RIP1, a member of an *Arabidopsis* protein family, interacts with the protein RARE1 and broadly affects RNA editing. Proc Natl Acad Sci U S A 109:E1453–E1461. https://doi.org/10.1073/pnas.1121465109

Bentolila S, Oh J, Hanson MR, Bukowski R (2013) Comprehensive high-resolution analysis of the role of an Arabidopsis gene family in RNA editing. PLoS Genet 9:e1003584. https://doi.org/10.1371/journal.pgen.1003584

Bisanz C, Bégot L, Carol P, Perez P, Bligny M, Pesey H, Gallois J-L, Lerbs-Mache S, Mache R (2003) The Arabidopsis nuclear *DAL* gene encodes a chloroplast protein which is required for the maturation of the plastid ribosomal RNAs and is essential for chloroplast differentiation. Plant Mol Biol 51:651–663

Blanc V, Litvak S, Araya A (1995) RNA editing in wheat mitochondria proceeds by a deamination mechanism. FEBS Lett 373:56–60. https://doi.org/10.1016/0014-5793(95)00991-H

Bock R (2000) Sense from nonsense: how the genetic information of chloroplastsis altered by RNA editing. Biochimie 82:549–557. https://doi.org/10.1016/S0300-9084(00)00610-6

Bock R, Koop HU (1997) Extraplastidic site-specific factors mediate RNA editing in chloroplasts. EMBO J 16:3282–3288. https://doi.org/10.1093/emboj/16.11.3282

Bock R, Kössel H, Maliga P (1994) Introduction of a heterologous editing site into the tobacco plastid genome: the lack of RNA editing leads to a mutant phenotype. EMBO J 13:4623–4628

Bock R, Hermann M, Kössel H (1996) *In vivo* dissection of *cis*-acting determinants for plastid RNA editing. EMBO J 15:5052–5059

Boussardon C, Salone V, Avon A, Berthomé R, Hammani K, Okuda K, Shikanai T, Small I, Lurin C (2012) Two interacting proteins are necessary for the editing of the NdhD-1 site in *Arabidopsis* plastids. Plant Cell 24:3684–3694. https://doi.org/10.1105/tpc.112.099507

Boussardon C, Avon A, Kindgren P, Bond CS, Challenor M, Lurin C, Small I (2014) The cytidine deaminase signature HxE(x)nCxxC of DYW1 binds zinc and is necessary for RNA editing of *ndhD-1*. New Phytol 203:1090–1095. https://doi.org/10.1111/nph.12928

Chateigner-Boutin A-L, Small I (2010) Plant RNA editing. RNA Biol 7:213–219

Chateigner-Boutin A-L, Ramos-Vega M, Guevara-García A, Andrés C, de la Luz Gutiérrez-Nava M, Cantero A, Delannoy E, Jiménez LF, Lurin C, Small I, León P (2008) CLB19, a pentatricopeptide repeat protein required for editing of *rpoA* and *clpP* chloroplast transcripts. Plant J 56:590–602. https://doi.org/10.1111/j.1365-313X.2008.03634.x

Chatterjee M, Sparvoli S, Edmunds C, Garosi P, Findlay K, Martin C (1996) DAG, a gene required for chloroplast differentiation and palisade development in *Antirrhinum majus*. EMBO J 15:4194–4207

Cheng S, Gutmann B, Zhong X, Ye Y, Fisher MF, Bai F, Castleden I, Song Y, Song B, Huang J, Liu X, Xu X, Lim BL, Bond CS, Yiu S-M, Small I (2016) Redefining the structural motifs that determine RNA binding and RNA editing by pentatricopeptide repeat proteins in land plants. Plant J 85:532–547. https://doi.org/10.1111/tpj.13121

Covello PS, Gray MW (1989) RNA editing in plant mitochondria. Nature 341:662–666. https://doi.org/10.1038/341662a0

Delannoy E, Stanley WA, Bond CS, Small ID (2007) Pentatricopeptide repeat (PPR) proteins as sequence-specificity factors in post-transcriptional processes in organelles. Biochem Soc Trans 35:1643–1647. https://doi.org/10.1042/BST0351643

Diaz MF, Bentolila S, Hayes ML, Hanson MR, Mulligan RM (2017) A protein with an unusually short PPR domain, MEF8, affects editing at over 60 Arabidopsis mitochondrial C targets of RNA editing. Plant J 92:638–649. https://doi.org/10.1111/tpj.13709

Doniwa Y, Ueda M, Ueta M, Wada A, Kadowaki K, Tsutsumi N (2010) The involvement of a PPR protein of the P subfamily in partial RNA editing of an *Arabidopsis* mitochondrial transcript. Gene 454:39–46. https://doi.org/10.1016/j.gene.2010.01.008

Farré JC, Araya A (2001) Gene expression in isolated plant mitochondria: high fidelity of transcription, splicing and editing of a transgene product in electroporated organelles. Nucleic Acids Res 29:2484–2491. https://doi.org/10.1093/nar/29.12.2484

Farré JC, Leon G, Jordana X, Araya A (2001) *cis* recognition elements in plant mitochondrion RNA editing. Mol Cell Biol 21:6731–6737. https://doi.org/10.1128/MCB.21.20.6731-6737.2001

Fujii S, Small I (2011) The evolution of RNA editing and pentatricopeptide repeat genes. New Phytol 191:37–47. https://doi.org/10.1111/j.1469-8137.2011.03746.x

García-Andrade J, Ramírez V, López A, Vera P (2013) Mediated plastid RNA editing in plant immunity. PLoS Pathog 9:1–13. https://doi.org/10.1371/journal.ppat.1003713

Giegé P, Brennicke A (1999) RNA editing in Arabidopsis mitochondria effects 441 C to U changes in ORFs. Proc Natl Acad Sci U S A 96:15324–15329. https://doi.org/10.1073/PNAS.96.26.15324

Glass F, Hä Rtel B, Zehrmann A, Verbitskiy D, Takenaka M (2015) MEF13 requires MORF3 and MORF8 for RNA editing at eight targets in mitochondrial mRNAs in *Arabidopsis thaliana*. Mol Plant 8:1466–1477. https://doi.org/10.1016/j.molp.2015.05.008

Gray MW (2012) Evolutionary origin of RNA editing. Biochemistry 51:5235–5242. https://doi.org/10.1021/bi300419r

Grewe F, Viehoever P, Weisshaar B, Knoop V (2009) A trans-splicing group I intron and tRNA-hyperediting in the mitochondrial genome of the lycophyte *Isoetes engelmannii*. Nucleic Acids Res 37:5093–5104. https://doi.org/10.1093/nar/gkp532

Grohmann L, Thieck O, Herz U, Schrüoder W, Brennicke A (1994) Translation of *nad9* mRNAs in mitochondria from *Solanum tuberosum* is restricted to completely edited transcripts. Nucleic Acids Res 22:3304–3311. https://doi.org/10.1093/nar/22.16.3304

Gualberto JM, Lamattina L, Bonnard G, Weil J-H, Grienenberger J-M (1989) RNA editing in wheat mitochondria results in the conservation of protein sequences. Nature 341:660–662. https://doi.org/10.1038/341660a0

Guillaumot D, Lopez-Obando M, Baudry K, Avon A, Rigaill G, Falcon De Longevialle A, Broche B, Takenaka M, Berthomé R, De Jaeger G, Delannoy E, Lurin C (2017) Two interacting PPR proteins are major Arabidopsis editing factors in plastid and mitochondria. Proc Natl Acad Sci U S A 114:8877–8882. https://doi.org/10.1073/pnas.1705780114

Gully BS, Shah KR, Lee M, Shearston K, Smith NM, Sadowska A, Blythe AJ, Bernath-Levin K, Stanley WA, Small ID, Bond CS (2015) The design and structural characterization of a synthetic pentatricopeptide repeat protein. Acta Crystallogr Sect D 71:196–208. https://doi.org/10.1107/S1399004714024869

Guo W, Grewe F, Fan W, Young GJ, Knoop V, Palmer JD, Mower JP (2016) *Ginkgo* and *Welwitschia* mitogenomes reveal extreme contrasts in gymnosperm mitochondrial evolution. Mol Biol Evol 33:1448–1460. https://doi.org/10.1093/molbev/msw024

Haag S, Schindler M, Berndt L, Brennicke A, Takenaka M, Weber G (2017) Crystal structures of the *Arabidopsis thaliana* organellar RNA editing factors MORF1 and MORF9. Nucleic Acids Res 45:4915–4928. https://doi.org/10.1093/nar/gkx099

Hackett JB, Shi X, Kobylarz AT, Lucas MK, Wessendorf RL, Hines KM, Bentolila S, Hanson MR, Lu Y (2017) An organelle RNA recognition motif protein is required for photosystem II subunit *psbF* transcript editing. Plant Physiol 173:2278–2293. https://doi.org/10.1104/pp.16.01623

Halter CP, Peeters NM, Hanson MR (2004) RNA editing in ribosome-less plastids of *iojap* maize. Curr Genet 45:331–337. https://doi.org/10.1007/s00294-003-0482-4

Hammani K, Okuda K, Tanz SK, Chateigner-Boutin A-L, Shikanai T, Small I (2009) A study of new *Arabidopsis* chloroplast RNA editing mutants reveals general features of editing factors and their target sites. Plant Cell 21:3686–3699. https://doi.org/10.1105/tpc.109.071472

Hayes ML, Hanson MR (2007) Assay of editing of exogenous RNAs in chloroplast extracts of *Arabidopsis*, maize, pea, and tobacco. Methods Enzymol 424:459–482. https://doi.org/10.1016/S0076-6879(07)24021-2

Hayes ML, Giang K, Berhane B, Mulligan RM (2013) Identification of two pentatricopeptide repeat genes required for RNA editing and zinc binding by C-terminal cytidine deaminase-like domains. J Biol Chem 288:36519–36529. https://doi.org/10.1074/jbc.M113.485755

Hayes ML, Dang KN, Diaz MF, Mulligan RM (2015) A conserved glutamate residue in the C-terminal deaminase domain of pentatricopeptide repeat proteins is required for RNA editing activity. J Biol Chem 290:10136–10142. https://doi.org/10.1074/jbc.M114.631630

He P, Huang S, Xiao G, Zhang Y, Yu J (2016) Abundant RNA editing sites of chloroplast protein-coding genes in *Ginkgo biloba* and an evolutionary pattern analysis. BMC Plant Biol 16:257. https://doi.org/10.1186/s12870-016-0944-8

He S, Sun Y, Yang Q, Zhang X, Huang Q, Zhao P, Sun M, Liu J, Qian W, Qin G, Gu H, Qu L-J (2017) A novel imprinted gene *NUWA* controls mitochondrial function in early seed development in *Arabidopsis*. PLoS Genet 13:e1006553. https://doi.org/10.1371/journal.pgen.1006553

Hegeman CE, Hayes ML, Hanson MR (2005) Substrate and cofactor requirements for RNA editing of chloroplast transcripts in Arabidopsis *in vitro*. Plant J 42:124–132. https://doi.org/10.1111/j.1365-313X.2005.02360.x

Herbert CJ, Golik P, Bonnefoy N (2013) Yeast PPR proteins, watchdogs of mitochondrial gene expression. RNA Biol 10:1477–1494. https://doi.org/10.4161/rna.25392

Hiesel R, Wissinger B, Schuster W, Brennicke A (1989) RNA editing in plant mitochondria. Science 246:1632–1634. https://doi.org/10.1126/SCIENCE.2480644

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. https://doi.org/10.1093/emboj/20.5.1144

Huang C, Yu Q-B, Li Z-R, Ye L-S, Xu L, Yang Z-N (2017) Porphobilinogen deaminase HEMC interacts with the PPR-protein AtECB2 for chloroplast RNA editing. Plant J 92:546–556. https://doi.org/10.1111/tpj.13672

Ichinose M, Sugita M (2017) RNA editing and its molecular mechanism in plant organelles. Genes (Basel) 8(1):5. https://doi.org/10.3390/genes8010005

Kim U, Wang Y, Sanford T, Zeng Y, Nishikura K (1994a) Molecular cloning of cDNA for double-stranded RNA adenosine deaminase, a candidate enzyme for nuclear RNA editing. Proc Natl Acad Sci U S A 91:11457–11461. https://doi.org/10.1073/pnas.91.24.11457

Kim U, Garner TL, Sanford T, Speicher D, Murray JM, Nishikura K (1994b) Purification and characterization of double-stranded RNA adenosine deaminase from bovine nuclear extracts. J Biol Chem 269:13480–13489

Kim S-R, Yang J-I, Moon S, Ryu C-H, An K, Kim K-M, Yim J, An G (2009) Rice *OGR1* encodes a pentatricopeptide repeat-DYW protein and is essential for RNA editing in mitochondria. Plant J 59:738–749. https://doi.org/10.1111/j.1365-313X.2009.03909.x

Kindgren P, Yap A, Bond CS, Small I (2015) Predictable alteration of sequence recognition by RNA editing factors from Arabidopsis. Plant Cell 27:403–416. https://doi.org/10.1105/tpc.114.134189

Knie N, Grewe F, Fischer S, Knoop V (2016) Reverse U-to-C editing exceeds C-to-U RNA editing in some ferns - a monilophyte-wide comparison of chloroplast and mitochondrial RNA editing suggests independent evolution of the two processes in both organelles. BMC Evol Biol 16:134. https://doi.org/10.1186/s12862-016-0707-z

Knoop V (2011) When you can't trust the DNA: RNA editing changes transcript sequences. Cell Mol Life Sci 68:567–586. https://doi.org/10.1007/s00018-010-0538-9

Knoop V, Rüdinger M (2010) DYW-type PPR proteins in a heterolobosean protist: plant RNA editing factors involved in an ancient horizontal gene transfer? FEBS Lett 584:4287–4291. https://doi.org/10.1016/j.febslet.2010.09.041

Kotera E, Tasaka M, Shikanai T (2005) A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. Nature 433:326–330. https://doi.org/10.1038/nature03229

Kugita M, Yamamoto Y, Fujikawa T, Matsumoto T, Yoshinaga K (2003) RNA editing in hornwort chloroplasts makes more than half the genes functional. Nucleic Acids Res 31:2417–2423

Lightowlers RN, Chrzanowska-Lightowlers ZM (2013) Human pentatricopeptide proteins: only a few and what do they do? RNA Biol 10:1433–1438. https://doi.org/10.4161/rna.24770

Lu B, Hanson MRR (1994) A single homogeneous form of ATP6 protein accumulates in petunia mitochondria despite the presence of differentially edited atp6 transcripts. Plant Cell 6:1955–1968. https://doi.org/10.1105/tpc.6.12.1955

Lu B, Wilson RK, Phreaner CG, Mulligan RM, Hanson MR (1996) Protein polymorphism generated by differential RNA editing of a plant mitochondrial *rps12* gene. Mol Cell Biol 16:1543–1549

Luo M, Cai M, Zhang J, Li Y, Zhang R, Song W, Zhang K, Xiao H, Yue B, Zheng Y, Zhao Y, Zhao J, Qiu F (2017) Functional divergence and origin of the *DAG-like* gene family in plants. Sci Rep 7:5688. https://doi.org/10.1038/s41598-017-05961-2

Lurin C, Andrés C, Aubourg S, Bellaoui M, Bitton F, Bruyère C, Caboche M, Debast C, Gualberto J, Hoffmann B, Lecharny A, Le Ret M, Martin-Magniette M-L, Mireau H, Peeters N, Renou J-P, Szurek B, Taconnat L, Small I (2004) Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. Plant Cell 16:2089–2103. https://doi.org/10.1105/tpc.104.022236

Lutz KA, Maliga P (2007) Transformation of the plastid genome to study RNA editing. Methods Enzymol 424:501–518. https://doi.org/10.1016/S0076-6879(07)24023-6

Mehta A, Driscoll DM (1998) A sequence-specific RNA-binding protein complements apobec-1 to edit apolipoprotein B mRNA. Mol Cell Biol 18:4426–4432. https://doi.org/10.1128/MCB.18.8.4426

Mehta A, Driscoll DM (2002) Identification of domains in apobec-1 complementation factor required for RNA binding and apolipoprotein-B mRNA editing. RNA 8:69–82

Melcher T, Maas S, Herb A, Sprengel R, Seeburg PH, Higuchi M (1996) A mammalian RNA editing enzyme. Nature 379:460–464. https://doi.org/10.1038/379460a0

Miyamoto T, Obokata J, Sugiura M (2002) Recognition of RNA editing sites is directed by unique proteins in chloroplasts: biochemical identification of *cis*-acting elements and *trans*-acting factors involved in RNA editing in tobacco and pea chloroplasts. Mol Cell Biol 22:6726–6734. https://doi.org/10.1128/MCB.22.19.6726-6734.2002

Naested H, Holm A, Jenkins T, Nielsen HB, Harris CA, Beale MH, Andersen M, Mant A, Scheller H, Camara B, Mattsson O, Mundy J (2004) *Arabidopsis VARIEGATED 3* encodes a chloroplast-targeted, zinc-finger protein required for chloroplast and palisade cell development. J Cell Sci 117:4807–4818. https://doi.org/10.1242/jcs.01360

Nakamura T, Sugita M (2008) A conserved DYW domain of the pentatricopeptide repeat protein possesses a novel endoribonuclease activity. FEBS Lett 582:4163–4168. https://doi.org/10.1016/j.febslet.2008.11.017

Neuwirt J, Takenaka M, van der Merwe JA, Brennicke A (2005) An in vitro RNA editing system from cauliflower mitochondria: editing site recognition parameters can vary in different plant species. RNA 11:1563–1570. https://doi.org/10.1261/rna.2740905.For

Oda K, Yamato K, Ohta E, Nakamura Y, Takemura M, Nozato N, Akashi K, Kanegae T, Ogura Y, Kohchi T, Ohyama K (1992) Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA. A primitive form of plant mitochondrial genome. J Mol Biol 223:1–7

Okuda K, Myouga F, Motohashi R, Shinozaki K, Shikanai T (2007) Conserved domain structure of pentatricopeptide repeat proteins involved in chloroplast RNA editing. Proc Natl Acad Sci 104:8178–8183. https://doi.org/10.1073/pnas.0700865104

Okuda K, Chateigner-Boutin A-L, Nakamura T, Delannoy E, Sugita M, Myouga F, Motohashi R, Shinozaki K, Small I, Shikanai T (2009) Pentatricopeptide repeat proteins with the DYW motif have distinct molecular functions in RNA editing and RNA cleavage in *Arabidopsis* chloroplasts. Plant Cell 21:146–156. https://doi.org/10.1105/tpc.108.064667

Okuda K, Hammani K, Tanz SK, Peng L, Fukao Y, Myouga F, Motohashi R, Shinozaki K, Small I, Shikanai T (2010) The pentatricopeptide repeat protein OTP82 is required for RNA editing of plastid ndhB and ndhG transcripts. Plant J 61:339–349. https://doi.org/10.1111/j.1365-313X.2009.04059.x

Okuda K, Shoki H, Arai M, Shikanai T, Small I, Nakamura T (2014) Quantitative analysis of motifs contributing to the interaction between PLS-subfamily members and their target RNA sequences in plastid RNA editing. Plant J 80:870–882. https://doi.org/10.1111/tpj.12687

Oldenkott B, Yamaguchi K, Tsuji-Tsukinoki S, Knie N, Knoop V (2014) Chloroplast RNA editing going extreme: more than 3400 events of C-to-U editing in the chloroplast transcriptome of the lycophyte *Selaginella uncinata*. RNA 20:1499–1506

Phreaner CGG, Williams MA, Mulligan RM (1996) Incomplete editing of *rps12* transcripts results in the synthesis of polymorphic polypeptides in plant mitochondria. Plant Cell 8:107–117. https://doi.org/10.1105/tpc.8.1.107

Pyo YJ, Kwon K-C, Kim A, Cho MH (2013) Seedling Lethal1, a pentatricopeptide repeat protein lacking an E/E+ or DYW domain in Arabidopsis, is involved in plastid gene expression and early chloroplast development. Plant Physiol 163:1844–1858. https://doi.org/10.1104/pp.113.227199

Rüdinger M, Funk HT, Rensing SA, Maier UG, Knoop V (2009) RNA editing: only eleven sites are present in the *Physcomitrella patens* mitochondrial transcriptome and a universal nomenclature proposal. Mol Gen Genomics 281:473–481. https://doi.org/10.1007/s00438-009-0424-z

Rudinger M, Fritz-Laylin L, Polsakiewicz M, Knoop V (2011) Plant-type mitochondrial RNA editing in the protist *Naegleria gruberi*. RNA 17:2058–2062. https://doi.org/10.1261/rna.02962911

Salone V, Rüdinger M, Polsakiewicz M, Hoffmann B, Groth-Malonek M, Szurek B, Small I, Knoop V, Lurin C (2007) A hypothesis on the identification of the editing enzyme in plant organelles. FEBS Lett 581:4132–4138

Schmitz-Linneweber C, Small I (2008) Pentatricopeptide repeat proteins: a socket set for organelle gene expression. Trends Plant Sci 13:663–670. https://doi.org/10.1016/j.tplants.2008.10.001

Shen C, Zhang D, Guan Z, Liu Y, Yang Z, Yang Y, Wang X, Wang Q, Zhang Q, Fan S, Zou T, Yin P (2016) Structural basis for specific single-stranded RNA recognition by designer pentatricopeptide repeat proteins. Nat Commun 7:11285. https://doi.org/10.1038/ncomms11285

Shi X, Hanson MR, Bentolila S (2015) Two RNA recognition motif-containing proteins are plant mitochondrial editing factors. Nucleic Acids Res 43:3814–3825. https://doi.org/10.1093/nar/gkv245

Shi X, Germain A, Hanson MR, Bentolila S (2016) RNA recognition motif-containing protein ORRM4 broadly affects mitochondrial RNA editing and impacts plant development and flowering. Plant Physiol 170:294–309. https://doi.org/10.1104/pp.15.01280

Shi X, Castandet B, Germain A, Hanson MR, Bentolila S (2017) ORRM5, an RNA recognition motif-containing protein, has a unique effect on mitochondrial RNA editing. J Exp Bot 68:2833–2847. https://doi.org/10.1093/jxb/erx139

Shikanai T (2006) RNA editing in plant organelles: machinery, physiological function and evolution. Cell Mol Life Sci 63:698–708. https://doi.org/10.1007/s00018-005-5449-9

Shikanai T (2015) RNA editing in plants: machinery and flexibility of site recognition. Biochim Biophys Acta 1847:779–785. https://doi.org/10.1016/j.bbabio.2014.12.010

Shikanai T, Fujii S (2013) Function of PPR proteins in plastid gene expression. RNA Biol 10:1446–1456. https://doi.org/10.4161/rna.25207

Sommer B, Köhler M, Sprengel R, Seeburg PH (1991) RNA editing in brain controls a determinant of ion flow in glutamate-gated channels. Cell 67:11–19. https://doi.org/10.1016/0092-8674(91)90568-J

Sosso D, Mbelo S, Vernoud V, Gendrot G, Dedieu A, Chambrier P, Dauzat M, Heurtevin L, Guyon V, Takenaka M, Rogowsky PM (2012) PPR2263, a DYW-subgroup pentatricopeptide repeat protein, is required for mitochondrial *nad5* and *cob* transcript editing, mitochondrion biogenesis, and maize growth. Plant Cell 24:676–691. https://doi.org/10.1105/tpc.111.091074

Staudinger M, Bolle N, Kempken F (2005) Mitochondrial electroporation and in organello RNA editing of chimeric *atp6* transcripts. Mol Gen Genomics 273:130–136. https://doi.org/10.1007/s00438-005-1117-x

Stein JC, Howlett B, Boyes DC, Nasrallah ME, Nasrallah JB (1991) Molecular cloning of a putative receptor protein kinase gene encoded at the self-incompatibility locus of *Brassica oleracea*. Proc Natl Acad Sci U S A 88:8816–8820

Stoltzfus A (2012) Constructive neutral evolution: exploring evolutionary theory's curious disconnect. Biol Direct 7:35. https://doi.org/10.1186/1745-6150-7-35

Sugita M, Ichinose M, Ide M, Sugita C (2013) Architecture of the PPR gene family in the moss *Physcomitrella patens*. RNA Biol 10:1439–1445

Sun T, Germain A, Giloteaux L, Hammani K, Barkan A, Hanson MR, Bentolila S (2013) An RNA recognition motif-containing protein is required for plastid RNA editing in Arabidopsis and maize. Proc Natl Acad Sci U S A 110:E1169–E1178. https://doi.org/10.1073/pnas.1220162110

Sun T, Shi X, Friso G, Van Wijk K, Bentolila S, Hanson MR (2015) A zinc finger motif-containing protein is essential for chloroplast RNA editing. PLoS Genet 11:e1005028. https://doi.org/10.1371/journal.pgen.1005028

Sun T, Bentolila S, Hanson MR (2016) The unexpected diversity of plant organelle RNA editosomes. Trends Plant Sci 21:962–973. https://doi.org/10.1016/j.tplants.2016.07.005

Sung T-Y, Tseng C-C, Hsieh M-H (2010) The SLO1 PPR protein is required for RNA editing at multiple sites with similar upstream sequences in Arabidopsis mitochondria. Plant J 63:1–13. https://doi.org/10.1111/j.1365-313X.2010.04258.x

Takenaka M (2010) MEF9, an E-subclass pentatricopeptide repeat protein, is required for an RNA editing event in the *nad7* transcript in mitochondria of Arabidopsis. Plant Physiol 152:939–947. https://doi.org/10.1104/pp.109.151175

Takenaka M (2014) How complex are the editosomes in plant organelles? Mol Plant 7:582–585. https://doi.org/10.1093/mp/sst170

Takenaka M, Brennicke A (2003) *In vitro* RNA editing in pea mitochondria requires NTP or dNTP, suggesting involvement of an RNA helicase. J Biol Chem 278:47526–47533. https://doi.org/10.1074/jbc.M305341200

Takenaka M, Verbitskiy D, van der Merwe JA, Zehrmann A, Plessmann U, Urlaub H, Brennicke A (2007) In vitro RNA editing in plant mitochondria does not require added energy. FEBS Lett 581:2743–2747. https://doi.org/10.1016/j.febslet.2007.05.025

Takenaka M, Zehrmann A, Verbitskiy D, Kugelmann M, Härtel B, Brennicke A (2012) Multiple organellar RNA editing factor (MORF) family proteins are required for RNA editing in mitochondria and plastids of plants. Proc Natl Acad Sci U S A 109:5104–5109. https://doi.org/10.1073/pnas.1202452109

Takenaka M, Zehrmann A, Verbitskiy D, Härtel B, Brennicke A (2013a) RNA editing in plants and its evolution. Annu Rev Genet 47:335–352. https://doi.org/10.1146/annurev-genet-111212-133519

Takenaka M, Zehrmann A, Brennicke A, Graichen K (2013b) Improved computational target site prediction for pentatricopeptide repeat RNA editing factors. PLoS One 8:e65343. https://doi.org/10.1371/journal.pone.0065343

Teng B, Burant CF, Davidson NO (1993) Molecular cloning of an apolipoprotein B messenger RNA editing protein. Science 260:1816–1819

Tillich M, Lehwark P, Morton BR, Maier UG (2006) The evolution of chloroplast RNA editing. Mol Biol Evol 23:1912–1921. https://doi.org/10.1093/molbev/msl054

Tillich M, Hardel SL, Kupsch C, Armbruster U, Delannoy E, Gualberto JM, Lehwark P, Leister D, Small ID, Schmitz-Linneweber C (2009) Chloroplast ribonucleoprotein CP31A is required for editing and stability of specific chloroplast mRNAs. Proc Natl Acad Sci U S A 106:6002–6007. https://doi.org/10.1073/pnas.0808529106

Verbitskiy D, Takenaka M, Neuwirt J, van der Merwe JA, Brennicke A (2006) Partially edited RNAs are intermediates of RNA editing in plant mitochondria. Plant J 47:408–416. https://doi.org/10.1111/j.1365-313X.2006.02794.x

Verbitskiy D, van der Merwe Ja, Zehrmann A, Brennicke A, Takenaka M (2008) Multiple specificity recognition motifs enhance plant mitochondrial RNA editing *in vitro*. J Biol Chem 283:24374–24381. https://doi.org/10.1074/jbc.M803292200

Verbitskiy D, Zehrmann A, Hartel B, Brennicke A, Takenaka M (2012) Two related RNA-editing proteins target the same sites in mitochondria of *Arabidopsis thaliana*. J Biol Chem 287:38064–38072. https://doi.org/10.1074/jbc.M112.397992

Wagner RW, Smith JE, Cooperman BS, Nishikura K (1989) A double-stranded RNA unwinding activity introduces structural alterations by means of adenosine to inosine conversions in mammalian cells. Proc Natl Acad Sci U S A 86:2647–2651. https://doi.org/10.1007/s12035-011-8220-2

Wagoner JA, Sun T, Lin L, Hanson MR (2015) Cytidine deaminase motifs within the DYW domain of two pentatricopeptide repeat-containing proteins are required for site-specific chloroplast RNA editing. J Biol Chem 290:2957–2968. https://doi.org/10.1074/jbc.M114.622084

Yagi Y, Hayashi S, Kobayashi K, Hirayama T, Nakamura T (2013) Elucidation of the RNA recognition code for pentatricopeptide repeat proteins involved in organelle RNA editing in plants. PLoS One 8:e57286. https://doi.org/10.1371/journal.pone.0057286

Yan J, Zhang Q, Guan Z, Wang Q, Li L, Ruan F, Lin R, Zou T, Yin P (2017) MORF9 increases the RNA-binding activity of PLS-type pentatricopeptide repeat protein in plastid RNA editing. Nat Plants 3:17037. https://doi.org/10.1038/nplants.2017.37

Yin P, Li Q, Yan C, Liu Y, Liu J, Yu F, Wang Z, Long J, He J, Wang H-W, Wang J, Zhu J-K, Shi Y, Yan N (2013) Structural basis for the modular recognition of single-stranded RNA by PPR proteins. Nature 504:168–171. https://doi.org/10.1038/nature12651

Zehrmann A, Verbitskiy D, van der Merwe JA, Brennicke A, Takenaka M (2009) A DYW domain-containing pentatricopeptide repeat protein is required for RNA editing at multiple sites in mitochondria of *Arabidopsis thaliana*. Plant Cell 21:558–567. https://doi.org/10.1105/tpc.108.064535

Zehrmann A, Verbitskiy D, Härtel B, Brennicke A, Takenaka M (2010) RNA editing competence of *trans*-factor MEF1 is modulated by ecotype-specific differences but requires the DYW domain. FEBS Lett 584:4181–4186. https://doi.org/10.1016/j.febslet.2010.08.049

Zehrmann A, Härtel B, Glass F, Bayer-Császár E, Obata T, Meyer E, Brennicke A, Takenaka M (2015) Selective homo- and heteromer interactions between the multiple organellar RNA editing factor (MORF) proteins in *Arabidopsis thaliana*. J Biol Chem 290:6445–6456. https://doi.org/10.1074/jbc.M114.602086

Zeltz P, Hess WR, Neckermann K, Börner T, Kössel H (1993) Editing of the chloroplast *rpoB* transcript is independent of chloroplast translation and shows different patterns in barley and maize. EMBO J 12:4291–4296

Zhang F, Tang W, Hedtke B, Zhong L, Liu L, Peng L, Lu C, Grimm B, Lin R (2014) Tetrapyrrole biosynthetic enzyme protoporphyrinogen IX oxidase 1 is required for plastid RNA editing. Proc Natl Acad Sci U S A 111:2023–2028. https://doi.org/10.1073/pnas.1316183111

Zhang H-D, Cui Y-L, Huang C, Yin Q-Q, Qin X-M, Xu T, He X-F, Zhang Y, Li Z-R, Yang Z-N (2015) PPR protein PDM1/SEL1 is involved in RNA editing and splicing of plastid genes in *Arabidopsis thaliana*. Photosynth Res 126:311–321. https://doi.org/10.1007/s11120-015-0171-4

Zhou W, Cheng Y, Yap A, Chateigner-Boutin A-L, Delannoy E, Hammani K, Small I, Huang J (2008) The Arabidopsis gene *YS1* encoding a DYW protein is required for editing of *rpoB* transcripts and the rapid development of chloroplasts during early growth. Plant J 58:1–15. https://doi.org/10.1111/j.1365-313X.2008.03766.x

Zhu Q, Dugardeyn J, Zhang C, Takenaka M, Kühn K, Craddock C, Smalle J, Karampelias M, Denecke J, Peters J, Gerats T, Brennicke A, Eastmond P, Meyer EHEH, Van Der Straeten D, Kuhn K, Craddock C, Smalle J, Karampelias M, Denecke J, Peters J, Gerats T, Brennicke A, Eastmond P, Meyer EHEH, Van Der SD (2012) SLO2, a mitochondrial pentatricopeptide repeat protein affecting several RNA editing sites, is required for energy metabolism. Plant J 71:836–849. https://doi.org/10.1111/j.1365-313X.2012.05036.x