Mitochondrial mRNA Processing in the Chlorophyte Alga *Pediastrum duplex* and Streptophyte Alga *Chara vulgaris* Reveals an Evolutionary Branch in Mitochondrial mRNA Processing

Grayson C. R. Proulex [1,†], **Marcus J. Meade** [1,†], **Kalina M. Manoylov** [2] and **A. Bruce Cahoon** [1,*]

1   Department of Natural Sciences, The University of Virginia's College at Wise, 1 College Ave., Wise, VA 24293, USA; gcp5a@uvawise.edu (G.C.R.P.); mjm6mv@uvawise.edu (M.J.M.)
2   Department of Biological and Environmental Sciences, Georgia College and State University, Milledgeville, GA 31061, USA; kalina.manoylov@gcsu.edu
*   Correspondence: abc6c@uvawise.edu; Tel.: +1-276-328-0201
†   These authors contributed equally to this project.

Abstract: Mitochondria carry the remnant of an ancestral bacterial chromosome and express those genes with a system separate and distinct from the nucleus. Mitochondrial genes are transcribed as poly-cistronic primary transcripts which are post-transcriptionally processed to create individual translationally competent mRNAs. Algae post-transcriptional processing has only been explored in *Chlamydomonas reinhardtii* (Class: Chlorophyceae) and the mature mRNAs are different than higher plants, having no 5′ UnTranslated Regions (UTRs), much shorter and more variable 3′ UTRs and polycytidylated mature mRNAs. In this study, we analyzed transcript termini using circular RT-PCR and PacBio Iso-Seq to survey the 3′ and 5′ UTRs and termini for two green algae, *Pediastrum duplex* (Class: Chlorophyceae) and *Chara vulgaris* (Class: Charophyceae). This enabled the comparison of processing in the chlorophyte and charophyte clades of green algae to determine if the differences in mitochondrial mRNA processing pre-date the invasion of land by embryophytes. We report that the 5′ mRNA termini and non-template 3′ termini additions in *P. duplex* resemble those of *C. reinhardtii*, suggesting a conservation of mRNA processing among the chlorophyceae. We also report that *C. vulgaris* mRNA UTRs are much longer than chlorophytic examples, lack polycytidylation, and are polyadenylated similar to embryophytes. This demonstrates that some mitochondrial mRNA processing events diverged with the split between chlorophytic and streptophytic algae.

Keywords: mitochondria; RNA processing; algal evolution; circular RNA; polycytidylation; PacBio Iso-Seq

1. Introduction

Mitochondria are membrane-bound organelles known for supplying eukaryotic cells with energy through ATP to carry out cellular functions. This occurs due to aerobic respiration whereby pyruvate is oxidized to CO_2 to generate reduced cofactors that drive the electron transport chain to chemiosmotically fuel ATP synthesis [1]. Despite the crucial role mitochondria have in supplying energy necessary for cellular functions and ATP for other biochemical pathways, it did not originate as a component of the eukaryotic cell. During the late 20th century, the theory of endosymbiosis became widely accepted and states that an aerobic bacterium was absorbed by, and formed an endosymbiotic relationship with, a pre-eukaryotic cell [2,3]. Though it became fully integrated into the Last Eukaryotic Common Ancestor (LECA), the proposed alpha-proteobacterium [4] maintained a portion of its circular genome carrying a conserved set of genes enabling the quick modulation of crucial energy acquisition proteins [5]. This remnant of the bacterial genome is referred to as mtDNA, the mitochondrial genome or chondriome.

Mitochondria have retained their own gene expression machinery, combining bacterial-like traits with novel features that evolved in the host cell [6]. Quite a bit is known about mitochondrial transcription and RNA processing from the compact chondriomes of humans and mice, which can serve for an overview of the process. Briefly, a nuclear-encoded RNA polymerase similar to those found in T3 and T7 bacteriophages [7] recognizes a promoter on both strands of mtDNA with the aid of two transcription factors [8,9]. These promoters occur in the only non-coding region (hyper-variable) and produce two long poly-cistronic primary RNAs known as heavy and light [10–14] with the aid of an elongation factor [14]. Individual mRNAs and tRNAs are removed from the primary transcripts by endonucleolytic cleavage by the enzymes RNaseP and RNaseZ, which precisely remove tRNAs, leaving most of the mRNAs as individual mRNAs with very short 3′ and 5′ UnTranslated Regions (UTRs), a process called the Punctuation Model [15–17]. Endonucleolytic processing between mRNAs with no intervening tRNA and between an mRNA with an adjacent antisense tRNA has been documented [13], but the enzymatic mechanisms responsible for these processing events are currently unknown. Having no 5′ UTRs, these mRNAs lack canonical ribosome binding sequences and use an alternative ribosome binding mechanism that is unique to mitochondria [18]. Once cleaved from the primary transcript, mRNAs may be polyadenylated, which adds the final adenine in some transcripts' stop codons, stabilizes some, and acts as a degradation signal for truncated messages [19–21]. mRNA fragments, but not full-length mRNAs, may also be circularized [22].

The mitochondrial genomes of plants (embryophytes) are much larger than those in animal cells due to expansive intergenic regions, repetitive DNA, and introns [23]. Plant mitochondria share some transcriptional processes with vertebrates. Transcription in plant mitochondria is catalyzed by one or more nuclear encoded phage-like RNA polymerases [24,25], and transcription factors similar to those used by vertebrates are encoded in plant nuclear genomes [26], but their functions have yet to be demonstrated. Due to their sizes, plant chondriomes have multiple promoters dispersed throughout the chondriome [27], yielding multiple primary poly-cistronic transcripts. Post-transcriptional processing takes on an expanded role in plants requiring numerous RNA Processing Factors (RPFs) that target endo- and exo-nucleolytic enzymes, define mRNA ends, and modify transcripts [28]. The 5′ termini of genes directly downstream of a transcriptional promoter are formed by the initial nucleotide added by the RNA polymerase [29]. For downstream genes in poly-cistronic transcripts, endonucleolytic cleavage between two genes will simultaneously produce the 5′ UTR of one gene and the 3′ UTR of an adjacent one. The lengths of these UTRs range from dozens to thousands of nucleotides consistent with the large intergenic regions of plant chondriomes [29–31]. To date, the best-defined cleavage mechanisms in plants are the precise removal of tRNAs by RNaseZ and PRORP. Similarly, tRNA-like secondary structures called t-elements also define intergenic cleavage sites recognized by endonucleases [31–35]. Most protein-coding genes are not separated by tRNAs, and their intergenic cleavage mechanism is hypothetical at this time but involves at least two nucleases [30,36]. Multiple 5′ termini for each gene usually result from these processes [30,31]. The 3′ ends are less variable and gene specific RPFs bind to them, presumably defining and stabilizing them [37–41]. The prevalence of group I and II introns in plant mitochondria creates an added layer of post-transcriptional processing. Neither class of intron is able to self-splice, so a group of nuclear-encoded RPFs are necessary for their removal [42]. In addition to the major construction of the mRNA coding regions, individual nucleotides are modified in a process known as RNA editing, which is common in higher plant mitochondria and chloroplasts [43]. Once the mRNA is no longer needed, it may be marked for degradation by way of polyadenylation by nuclear encoded factors in a manner similar to that of bacteria [44–47].

Our understanding of mitochondrial transcription and RNA processing in algal species is mostly limited to the single-celled photosynthetic green alga *Chlamydomonas reinhardtii* P.A. Dangeard, which is a well-established model system [48]. *C. reinhardtii* has a

small linear chondriome [49] that is unusual among algae but is a conserved trait among the Reinhardtinia clade of the Order Chlamydomonadales [50,51]. In this species, transcription is initiated on each of the two strands from promoters in a small intergenic region to produce two primary transcripts [52,53]. Each mRNA is endonucleolytically cleaved directly adjacent to the AUG start codon, leaving no 5′ UTR, similar to those seen in animal systems. The 3′ UTRs are comprised of various lengths of template-derived intergenic regions and may have non-template polycytosine and/or polyuracil tails added, presumably as part of the maturation process [54–56]. The poly-cytidylation of mitochondrial mRNAs seen in green algae is unusual and appears to be limited to the algal class Chlorophyceae [55]. It has been hypothesized that these leaderless mRNAs use an alternative ribosome-binding mechanism, but there is evidence that the mature mRNAs are circularized, which brings putative ribosome binding sites (RBSs) located in the intergenic regions of the *Chlamydomonas* chondriome upstream of the start codon to initiate translation [56]. mRNAs in *C. reinhardtii* are also poly-adenylated, which serves as a degradation signal consistent with mitochondria in other eukaryotes and bacteria [54,57,58]. mRNA editing, which is common among embryophytes, is missing in both the chlorophyte and streptophyte lineages of green algae [59], suggesting that some post-transcriptional processes were acquired by embryophytes after they invaded land.

The purpose of this study was to define the 5′ and 3′ UTRs of mitochondrial mRNAs in two algae, *Pediastrum duplex* Meyen (P: Chlorophyta, C: Chlorophyceae, F: Hydrodictyaceae) from the chlorophyte algal clade and *Chara vulgaris* Linnaeus (P: Charophyta, C: Charophyceae, F: Characeae) of the charophyte algal clade. *P. duplex* is a member of the same class as *C. reinhardtii*, but it has a circular chondriome that is several times larger [60,61], making the architecture more similar to that found in other algae. By defining the termini, we hoped to determine if RNA processing events seen in *C. reinhardtii* are also used among other chlorophytic algae or are related to its compact genome. *C. vulgaris* has a mitochondrial genome similar in size to *P. duplex*, but the gene content and synteny are more similar to bryophytes [62,63]. We analyzed the 3′ and 5′ ends from *C. vulgaris* to see if mRNA end processing resembled higher plants or chlorophytic green algae. Circular RT-PCR (cRT-PCR) and PacBio long-read sequencing were used to define the transcript termini of 12 mitochondrial mRNAs from each species. cRT-PCR allows the mapping of both the 3′ and 5′ ends of an RNA by artificially ligating them together followed by the production of a cDNA across the ligation site, PCR amplification of the sequences flanking the ligation site and sequencing of those amplicons. PacBio Iso-Seq is a long-read RNA sequencing technology which can sequence full-length RNAs including their 3′ and 5′ termini. This platform sequences poly-adenylated RNAs so organellar mRNAs must be artificially poly-adenylated to increase the likelihood they will be sequenced.

2. Results

2.1. Pediastrum duplex

In *P. duplex*, *cox2a*, *cox3*, *nad2*, *nad4*, and *nad4L* 5′ termini occurred directly upstream of the AUG start codon, essentially leaving no UTR (Table 1). For *atp6*, 5′ and 3′ end processing occurred within a 9 nt genomic DNA (gDNA) encoded stretch of adenines flanking the gene (Figure 1A). Since the same oligonucleotide sequence occurred at both ends, it was not possible to distinguish the exact location of the 5′ or 3′ exonucleolytic cleavage using cRT-PCR. Two of the *P. duplex* mRNAs (*atp9* and *nad1*) had cleavage sites producing 5′ UTR termini downstream of the predicted start codons in archived chondriome maps (KR026340, KR026340, MK895949). For *atp9* (Figure 1B) and *nad1* (Figure 1D), there is an in-frame AUG start codon adjacent to the cut site leaving a short 5′ UTR consistent with the other genes. For *cob*, there was disagreement in the 5′ terminus between the circRT-PCR and PacBio techniques. Using circ-RT-PCR, a single 5′ terminus 22 nt downstream of the predicted start codon was detected (Figure 1C), while reads using PacBio IsoSeq revealed a single terminus adjacent to the originally predicted start codon.

Table 1. Site of 5′ UTR terminus upstream from start codon (in nucleotides) in *Pediastrum duplex* mitochondria, nd = no data. * Some 5′ UTR termini detected in this study occur downstream of the start codons in archived chondriomes. The distances presented in this table are marked from the next available AUG start codon.

	Site of 5′ UTR Terminus Upstream from Start Codon (in Nucleotides)	
Gene	Circular RT-PCR	PacBio
atp6	−2–11	nd
atp9	−3 *	nd
cob	−2 *	0
cox1	nd	0
cox2a	0	nd
cox3	0	0
nad1	−15 *	nd
nad2	0	nd
nad4	0	0
nad4L	0	0
	−81–93	nd
nad5	nd	nd
nad6	nd	0
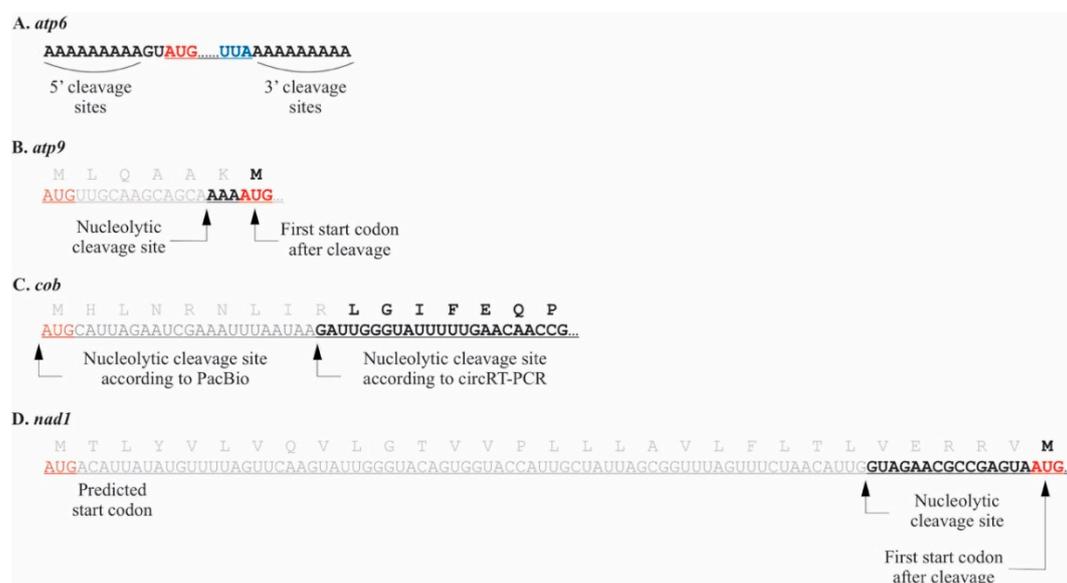
Figure 1. The 5′ termini and UTRs of four *P. duplex* mitochondrial mRNAs. Start codons are red, stop codons blue, and truncated portions of coding regions grey. (**A**). The 5′ UTR of the *atp6* mRNA occurred within a 9 nt templated stretch of adenines that also appears downstream of the stop codon. (**B**). The *atp9* 5′ terminus occurred downstream of the AUG start codon in archived *P. duplex* chondriomes, suggesting the protein may be six amino acids shorter than predicted. (**C**). Two *cob* 5′ termini were detected. Circular RT-PCR revealed one downstream of the AUG start codon in archived *P. duplex* chondriomes, while the PacBio Iso-Seq technique revealed a terminus directly adjacent to the predicted start codon. (**D**). The *nad1* 5′ terminus occurred downstream of the AUG start codon in archived *P. duplex* chondriomes, suggesting the protein may be thirty amino acids shorter than predicted.

P. duplex 3′ UTR lengths were gene specific, with several having two or more termini (Figure 2A–J and Table 2). Most were relatively short, fewer than 25 nts, the exception being *cob* which were 100 and 110 nts in length (Figure 2C). Eight genes (*atp6*, *atp9*, *cob*, *cox1*, *cox2a*, *cox3*, *nad2*, and *nad5*) were polycytidylated. For some, this occurred at specific

termini, e.g., *atp6*, *atp9*, *cob*, *cox1*, *cox2a*, and *cox3*. For two genes, *nad2* and *nad5*, the poly(C) additions occurred at variable locations within templated repetitive AU regions beginning 9nts downstream of the stop codons (Figure 2G,J and Table 2). No poly(C) additions were detected on *nad1*, *nad4*, or *nad4L*. On *nad4L*, there was an AU repeat region adjacent to the stop codon, but no poly(C) additions were detected. There was general agreement of the 3′ termini of fully processed mRNAs between cRT-PCR and PacBio reads for all but one transcript, *nad4L*. For this gene, cRT-PCR provided one 3′ terminus with a truncated stop codon (Figure 2J upper sequence), whereas PacBio data provided a 3′ terminus 37 nt downstream (Figure 2J lower sequence). The nucleotide sequences between the stop codon and the 3′ terminus for each gene were aligned and analyzed using a logo plot, and the 15 nucleotides upstream of the terminus were comprised nearly exclusively of adenines and uracils (Supplemental Figure S1A). These regions were analyzed for RNA secondary structure and none were detected.

Naturally circularized mRNAs were also detected using RT-PCR. In *P. duplex*, circularized variants carrying full-length coding regions were detected for seven mRNAs (*atp6*, *atp9*, *cob*, *cox2a*, *cox3*, *nad4L*, and *nad5*) (Figure 3). For five of these, the circularization coincided with a tandemly repeated nucleotide motif. The ligation site for naturally circularized *atp6* transcripts occurred within a stretch of template coded adenines (Figure 3A). Several circularized variants of *cob* were found (Figure 3C), one where the circularization occurred within a template encoded polyU motif and a second within two AU rich motifs. Two *cox3* circular transcript variants were detected, one where the ligation occurred within a GAACGAA motif and a second ligated at a GCGTCTT motif that removed the final 45 nts of the coding region. Two *nad4L* circular variants were detected, each occurring within AT rich motifs and including full-length coding regions. Three naturally circularized transcripts were detected with poly(C) additions. The two circular variants of *atp6* and *cob* with a poly(C) addition had severely truncated coding regions (Figure 3A,C). Two *atp9* circular variants were detected and both had poly(C) additions and full-length coding regions (Figure 3B). A single circularized variant of the *cox2a* transcript was detected, with no obvious repeat motif and no poly(C) additions (Figure 3D).

The long-read PacBio data did not cover the entire chondriome or contain reads longer than ~2200 nt but, when combined with the circ-RT-PCR data, they did allow the detection of some broader transcript processing events from three portions of the *P. duplex* chondriome (Figure 4). For *cox1*, reads spanning the two exons and the intron were detected (Figure 4A). All of these reads had 5′ termini directly adjacent to the start AUG, while some were polycytidylated on the 3′ terminus. Transcripts with the intron removed were also detected that had the same 5′ and 3′ termini as the unspliced transcript. PacBio reads were also produced for another section with three genes, *nad2-nad6-cob*, flanked by tRNAs (Figure 4B). For *nad2*, transcripts appeared to have been endonucleolytically cleaved adjacent to *trnN*, forming a 5′ terminus for *nad2* and adjacent to *nad6* to form a 3′ terminus. The 5′ end was further processed, leaving multiple termini eventually resulting in an mRNA with no UTR that was also polycytidylated on the 3′ terminus. A transcript with both *nad6* and *cob* was detected. Its 5′ terminus occurred adjacent to *nad6*'s start codon, while the 3′ terminus appeared to have been created by the cleavage of *trnV* from the primary transcript followed by polycytidylation. The *cob* coding region was cleaved away from *nad6*, leaving two different 5′ termini. The linear transcripts were polycytidylated, whereas the circular version of *cob* was ligated with no poly(C) tract. No individual *nad6* transcript was detected from either circRT-PCR or PacBio results. A third region with the two genes *nad4L-atp9* flanked by tRNAs was also produced by the PacBio Iso-Seq methodology (Figure 4C). A single transcript that appeared to have been produced by the removal of the two tRNAs from the primary transcript was detected. The 3′ terminus of this poly-cistronic mRNA was poly-citidylated after *trnE* was removed. The removal of the *trnG* occurring 5′ of *nad4L* left 170 nt upstream of the start codon, but this was removed, leaving a 5′ terminus adjacent to the start AUG. The 3′ terminus was produced by endonucleolytic cleavage, leaving a 37 nt 3′ UTR that was polycytidylated. A second *nad4L* transcript was

detected with a longer 5′ UTR (−80–94 nt within an AU repeat region) and a 3′ terminus comprised of a truncated stop codon that had been polycytidylated. Both versions of these shortest *nad4L* transcripts were detected as circular RNAs and neither contained a poly(C) tract. The *atp9* coding region was cleaved from the primary transcript, leaving a 5′ terminus −2 nt upstream of its start AUG, but its 3′ terminus was the one formed by the removal of *trnE*. A non-polycytidylated version of this mRNA was circularized.

A. *atp6*   …**UUA**AAAAAAAA
              …**UUA**AAAAAAAAGUAUAUAUAAAA-polyC

B. *atp9*   …**UAA**UAAAAAAAUAAAUAAAUAAAUAA-polyC
              …**UAA**UAAAAAAAUAAAUAAAUAAAUAAUUU

C. *cob*   …**UAA**CAAGAUAGGUCUAUAGGAUAUAAGUAUAGAUUUGCGUCUUU
              UCCAAUUCUAUUUUUACCUCUUUAUAGGUCAUAGAUUUAAACAUA
              GUUUUUAAUUAUUU-polyC

              …**UAA**CAAGAUAGGUCUAUAGGAUAUAAGUAUAGAUUUGCGUCUUU
              UCCAAUUCUAUUUUUACCUCUUUAUAGGUCAUAGAUUUAAACAUA
              GUUUUUAAUUAUUUAUAUAUAAUA

D. *cox2a*   …**UAA**AAAUUUUUUU-polyC

E. *cox3*   …**UAA**UGUUUAACUCUUA-polyC

F. *nad1*   …**UCA**ACUAAUUAUUUUUAAAUU

G. *nad2*   …**UAA**AACAUUAAAUAUAUAUAUAUA
                                  ⌐ polyC additions occurred within
                                  ⌐ this AU repeat region

H. *nad4*   …**UAA**UUUUUUUUA

I. *nad4L*   …**UAA**
              …**UUA**UAUAUAUAUAUAUAUAUAGUCAUAAAUGACUAGAUAU

J. *nad5*   …**UUA**AUUUUUUUUCUAUAUAUAUAUAUAUAU
                                  ⌐ polyC additions occurred within
                                  ⌐ this AU repeat region

Figure 2. The 3′ termini and UTRs of ten *P. duplex* mitochondrial mRNAs. Stop codons are blue. (**A**). Two 3′ termini were detected for *atp6*. The upper (shorter) sequence had no oligonucleotide addition, while a portion of transcripts represented by the lower (longer) sequence had poly(C) additions. (**B**). For *atp9* a portion of the transcripts with the upper terminus had a poly(C) addition, while the slightly longer one represented by the lower sequence did not. (**C**). For *cob* a portion of the transcripts with the upper terminus had a poly(C) addition, while the one represented by the longer sequence did not. (**D**). For *cox2a* a single terminus was detected and a portion of them had a poly(C) tail. (**E**). A single terminus, some of which had a poly(C) addition, was detected for *cox3*. (**F**). A single terminus was detected for *nad1* and no oligonucleotide additions were detected. (**G**). For *nad2* poly(C) additions were detected at several different termini within an AU repeat region. (**H**). A single terminus was detected for *nad4* and no oligonucleotide additions were detected. (**I**). A single terminus was detected for *nad4L* and no oligonucleotide additions were detected. (**J**). For *nad5* poly(C) additions were detected at several different termini within an AU repeat region.

Table 2. Lengths of the most abundant 3′ UTRs found in *Pediastrum duplex* and those found with a non-template addition

Gene	Circular RT-PCR				PacBio Long Reads		Genes with an Adjacent tRNA (Distance from Stop Codon in nt)
	Length of 3′ UTR from Stop Codon	UTRs with polyC Addition	UTRs with polyA + polyC Addition	UTR with polyA Addition	Length of 3′ UTR from Stop Codon	UTRs with polyC Addition	
atp6	21	+	-	-	nd	nd	
	0–9	-	-	-	nd	nd	
atp9	22	+	-	-	16–22	+	26
	25	-	-	-	nd	nd	
cob	110	-	-	-	104–110	+	110
	100	+	-	-	nd	nd	
cox1	13	+	-	-	7–14	+	
cox2a	10	+	-	-	nd	nd	28
cox3	13	+	-	-	10–13	-	
nad1	18	-	-	-	nd	nd	
nad2	9–21	+	-	-	5–23	-	
nad4	8	-	-	-	7–8	+	786
nad4L	−1	-	-	-	37	-	
nad5	0–8	-	-	+	1–3	-	42
	9–24	+	+				
nad6	nd	nd	nd	nd	nd	nd	

nd: no data.

Since tRNA removal was found to be integral in the maturation of *cob* and *atp9* 3′ termini, the placement of tRNAs was compared to the 3′ termini of other genes (Table 2). Five of the genes we analyzed had a 3′ adjacent tRNA, but only two, *cob* and *atp9*, had mature 3′ ends matching the placement of those tRNAs. The possibility of t-elements was considered for the other genes, but no evidence of secondary structures immediately downstream of the mature 3′ ends was detected for *atp6*, *cox1*, *cox2a*, *cox3*, *nad1*, *nad2*, *nad4*, *nad4L*, *nad5*, or *nad6*.
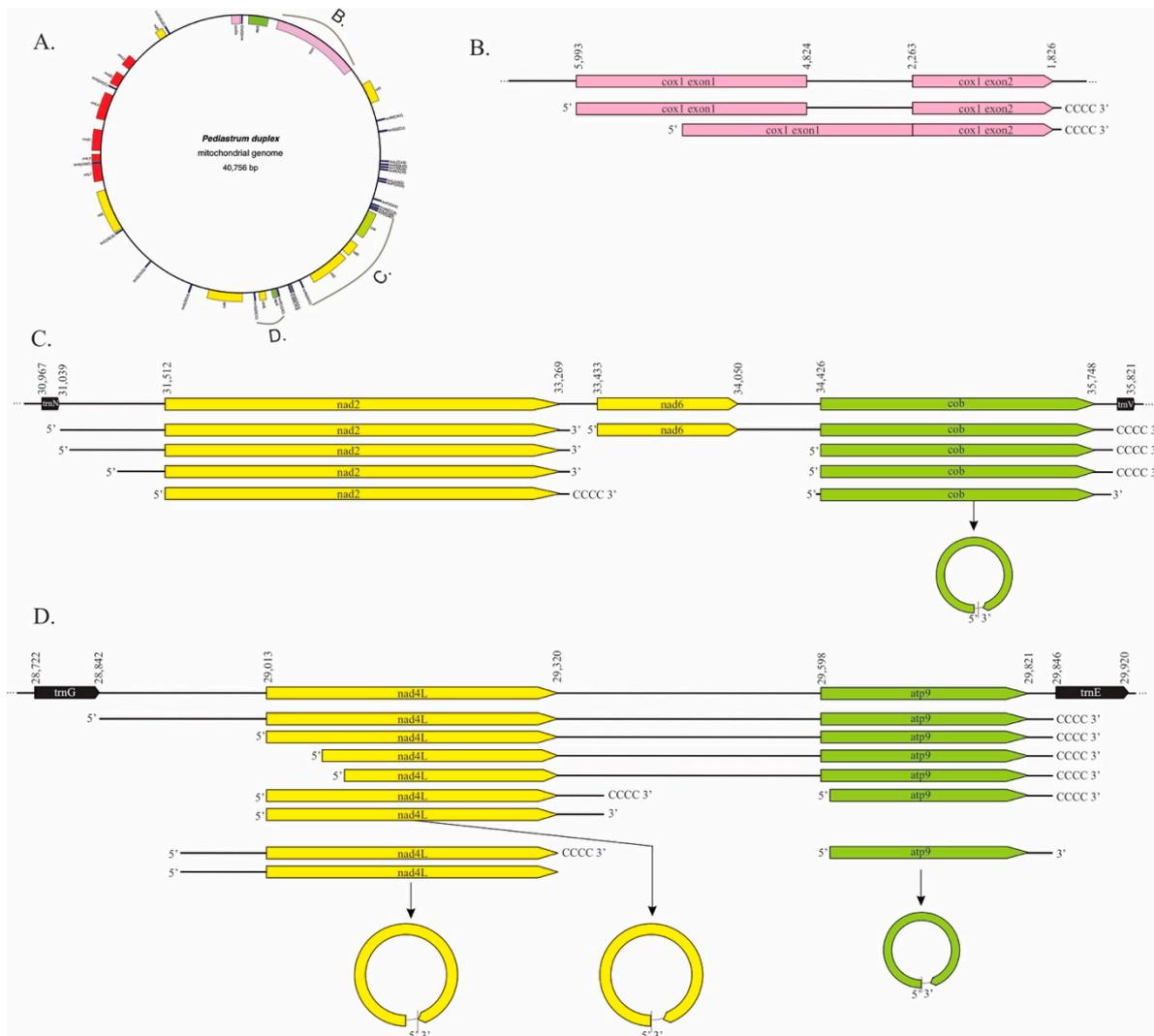
A. *atp6*　　**AAAAAAAGU**<u>**AUG**</u>...<u>**UAA**</u>**AAAAAAAA**

AUGACCAAUUUUAUAGCAUUCGUAUUUAACUCUUCUGCGGAAUAGUGGCAAAGUAUUCCAUUGAUUAGUUUUCAUUAUAGUUAGUUUAGUUUUAUU
UUUAGCGCAAAU**U**...<u>**UAA**</u>**AAAAAAAGUAUAUAUAAAA**-polyC

B. *atp9*　　M　L　Q　A　A　K　M
AUGUUGCAAGCAGC**AAAAAUG**...<u>**UAA**</u>UAAAAAAAUAAAUAAAUAA-polyC

AUGUUGCAAGCA**GCAAAAAUG**...<u>**UAA**</u>UAAAAAAAUAAAUAAAUAAAUAA-polyC

C. *cob*　　**GUUUUU**UUGUUUUUU<u>**AUG**</u>...<u>**UAA**</u>CAAGAUAGGUCUAUAGGAUAUAAGUAUAGAUUUGCGUCUUUUCCAAUUCUAUUUUUACCUCUUUAUAGGUCAUA
GAUUUAAACAUA**GUUUUU**

AUUAUUUUUAUAAAUAAGUUUUUUUGUUUUUU<u>**AUG**</u>...<u>**UAA**</u>CAAGAUAGGUCUAUAGGAUAUAAGUAUAGAUUUGCGUCUUUUCCAAUUCUAUUUUUA
CCUCUUUAUAGGUCAUAGAUUUAAACAUAGUUUUUA**AUUAUUUAUAUAUAAUA**

<u>AUG</u>CAUUAGAAUCGAAAUUUAAUAAGAUUGGGUAUUUUUGAACAACCGUUAUUAAAUGUAGUACAAAAUCAUCUAAUUGCUUAUCC**A**...
CAAGAUAGGUCUAUAGGAUAUAAGUAUAGAUUUGCGUCUUUUCCAAUUCUAUUUUUACCUCUUUAUAGGUCAUAGAUUUAAACAUAGUUUUUAAUU
<u>AUUUAUAUAUAAUA</u>
　　　　　　　└── polyA additions occurred
　　　　　　　　　within this region

D. *cox2a*　CGAAGAAGAAGCCGAUGCGCAUCGGCUUCUUCGCUUCGCGAACGAAGAAGCCAGCCGAAGCGAAAGAAAGCAGGGUGCGGGGGUGCACUUUGCUGU
CGCGGAGCGAAUGCAAACAAGAAAUAUUGAGCUUCUUAUAAAGCUGGGUAUUUAUUAUU<u>**AUG**</u>...<u>**UAA**</u>AAAUUUUUUUAUAUAUUAAAUUAUUU<u>GAGU</u>
<u>GUGUAGCUCAAUUGGUUAGAGCGCAGGACCGAUAAUCCUGAGGUUGCAGGUUCAAUCCCUGCCAUACUCAA</u>ACAAUGGCGUUAGCACUGGCUUCUU
<u>UCGCUUCGGCUGGCUUCUUCGUUCGCUUGCGCUUUCG</u>

AGCCGAUGCGCAUCGGCUUCUUCGCUUCGCGAACGAAGAAGCCAGCCGAAGCGAAAGAAAGCAGGGUGCGGGGGUGCACUUUGCUGUCGCGGAGCG
AAUGCAAACAAGAAAUAUUGAGCUUCUUAUAAAGCUGGGUAUUUAUUAUU<u>**AUG**</u>...<u>**UAA**</u>AAAUUUUUUUAU

E. *cox3*　**GAACGAA**GAAGCACCAGUUCUACGAAUCGCGUCUUGUCAGACGCGGGCAAGUCCUACGGACCUUUUAAAUUUUUUUUUUGUUUUUU<u>**AUG**</u>...<u>**UAA**</u>UGU
UUAACUCUUAUUUUUAUUUAUAAAAGUAGAGUAGUAUAAAAGUGGAGCCUAUAUAUAUAUAAUAUAUAAGUCAAAAACUAGAGUACUAGUAGUACU
CUAACUUUUCCACUUUGCUGUGUAUCUAAAAUUUUGAAAUGGGUACAGUGAGCUUAUUUAUUUUUUUUUUAUAUAUUUAUUUUUGUUUUUAUUGAC<u>G</u>
<u>UUAUAUUUGGAGCACCCGAAAGCAGGUGAUCUUCCCUUGUCCAAGUUGAAGGCUGGCUUUUUGCACCGCUGGAGGACUGAACUCGUGUACGUGGUA</u>
<u>AAGUACUGAGGUGAGGUGGGGGAAGGAGUGAAAGGCUAAUCAAACUUGCCUAUCGCUGG</u>UUUUU**GAACGAA**

GCGUCUUGUCAGACGCGGGCAAGUCCUACGGACCUUUUAAAUUUUUUUUUUGUUUUU<u>**AUG**</u>...**GCGUCUU**GAUUAGUAUAGUCAUAAUCCUGAAUUU
UUUGAAGCAUCUAUGUAA

F. *nad4L*　**UAUAUAUAUAUAUA**CAAAAUUUACAACUAAAUUUAUCUUUAUAUUAAUUAAAAGAUAAAUGUGUUUAUCUGUACUUUUAUUUAUAAAAGUACAG**AU**
G...<u>**UAA**</u>**UAUAUAUAUAUAUAUAUA**

UUUAUAAAAGUACAG**AUG**...<u>**UAA**</u>UAUAUAUAUAUAUAUAUAUAGUCAUAAAUGACUAGAUAUAAAACUGCA**UUUAUAAA**

G. *nad5*　**UAUAUAUG**...<u>**UAA**</u>UUUUUUUUUC**UAUAUAUAUAUAUAUAUA**

Figure 3. Naturally circularized mRNAs found in *P. duplex* mitochondria. Start codons are colored red, stop codons blue, repeat sequences orange, and truncated portions of coding regions grey. (**A**). Two circularized mRNAs were detected for the gene *atp6*, the upper sequence represents a full-length coding region whose 3′ and 5′ termini were ligated within templated adenine stretches that flank the coding region. The lower sequence represents a circularized portion of the mRNA. (**B**). Two versions of a circularized *atp9* transcript were detected. Both would be full length considering a start codon downstream of the previously predicted one (grey) would be the actual start codon. (**C**). Two full-length circularized versions of the *cob* transcript were detected (upper two sequences). The ligation termini coincided with two different repeat sequences (orange). A third circularized transcript with a truncated coding region was also detected for *cob*. (**D**). Two circularized full-length coding regions were detected for *cox2a*. (**E**). Two circularized *cox3* transcripts were detected, both ligations having occurred at repeat sequences. One carried a full-length coding region (upper) the other truncated (lower). (**F**). Two full-length circularized versions of the *nad4L* transcript were detected where ligation occurred at AU-rich repeat sequences. (**G**). One circularized *nad5* transcript was detected with the ligation occurring within AU-rich repeat sequences.

Figure 4. Mitochondrial RNA processing as determined by PacBio Iso-Seq and circRT-PCR. (**A**). The mitochondrial genome of *P. duplex* (GenBank MK895949). The regions highlighted in portions B, C, and D of this figure are marked. The map was generated using OGDRAW [64]. (**B**). The *cox1* gene, the upper diagram represents a portion of the chondriome, the lower two portions represent partially and fully processed transcripts. (**C**). The *trnG-nad2-nad6-cob-trnE* portion of the chondriome. (**D**). The *trnG-nad4L-atp9-trnE* portion of the chondriome.

2.2. Chara vulgaris

The 5′ UTRs were much longer than those observed in *P. duplex* (Table 3). Based on the termini we detected, UTRs ranged from 6–273 nucleotides, with an average length of 80 nts (S.E. = 15 nt). They were also more variable, with two or more termini detected for seven genes, *atp6*, *cob*, *cox2*, *cox3*, *nad1*, *nad4*, and *nad4L*. For three of the genes, *cox1*, *nad1*, and *nad2*, the mapped 5′ termini occurred downstream of the start codons in Gen-Bank record NC_005255 (Figure 5). For *cox1*, a single terminus was detected 125–129 nts downstream of the predicted start codon. The next start AUG occurs 75–79 nt downstream (Figure 5A). Three 5′ UTRs were detected for *nad1*, all of which remove the predicted start codon but, depending upon the cleavage site, leave two possible alternative AUG start codons (Figure 5B). In *nad2*, two 5′ UTRs were detected, both of which leave a single alternative AUG start codon (Figure 5C). The length of the 5′ UTRs in *C. vulgaris* raised the possibility that they may fold to form RNA secondary structures, but the probability of secondary structures in the 5′ UTRs was found to be extremely low. The coverage of chondriome derived transcripts for *C. vulgaris* using PacBio sequencing was very low,

so the 5′ termini of only two genes was recovered, *atp9* and *cox2*. Neither were the same length as those detected by circRT-PCR (Table 3).

Table 3. Site of 5′ UTR terminus upstream from start codon (in nucleotides) in *Chara vulgaris* mitochondria, nd = no data. * 5′ UTR termini detected in this study occur downstream of the start codons found in archived chondriomes. The distances presented in this table are marked from the next available AUG start codon.

	Site of 5′ UTR Terminus Upstream from Start Codon (in Nucleotides)	
Gene	**Circular RT-PCR**	**PacBio**
atp6	−76 −62	nd
atp9	−56	−93
cob	−161 −46	nd
cox1	−150 *	nd
cox2	−318 −18	−52
cox3	−153 −31	nd
nad1	−73 * −27 * −8 *	nd
nad2	−112 *	nd
nad4	−273 −107 −99 −82 −18 −6	nd
nad4L	−260 −18	nd
nad5	nd	nd
nad6	−36	nd

    C. vulgaris 3′ ends ranged from 0–162 nts with an average of 61.3, S.E. = 7.5, (Table 4, Figure 6). Multiple 3′ termini were detected for each gene, except *cox1*, and *nad2*, which had single termini. Polyadenylation was detected on eight genes (*atp6*, *atp9*, *cob*, *cox3*, *nad1*, *nad2*, *nad4*, and *nad4L*). For genes where multiple termini were detected, the polyA tail only occurred on one of those termini (Figure 6B,C,F,G,I). The exception was *atp6*, where all three termini were polyadenylated (Figure 6A). The proportion of those specific transcripts with polyA additions varied considerably. For example, 0.7% of *atp9* transcripts with 52–57 nt 3′ UTRs had a polyA tail, whereas the majority of specific *nad2*, *nad4*, and *nad4L* transcripts were tailed. PacBio sequencing produced data for four genes and the 3′ termini agreed with three (Table 4). The exception was *cob*, where the PacBio sequencing revealed a longer 3′ UTR than those found with circRT-PCR. PacBio sequencing could not be used to detect non-template poly(A) tails since the mitochondrial transcripts were artificially polyadenylated to accommodate the Iso-Seq technique. The forty nucleotides upstream of the 3′ termini were analyzed for conserved sequences using a logo plot and none were found (Supplemental Figure S1B). The length of the 3′ UTRs in *C. vulgaris* raised the possibility that these regions could fold into secondary structures. Secondary structure prediction suggested a high probability that stable stem-loop structures occur adjacent to the 3′ terminus in all but one (*cox2*) of these 3′ UTRs (Supplemental Figure S2).
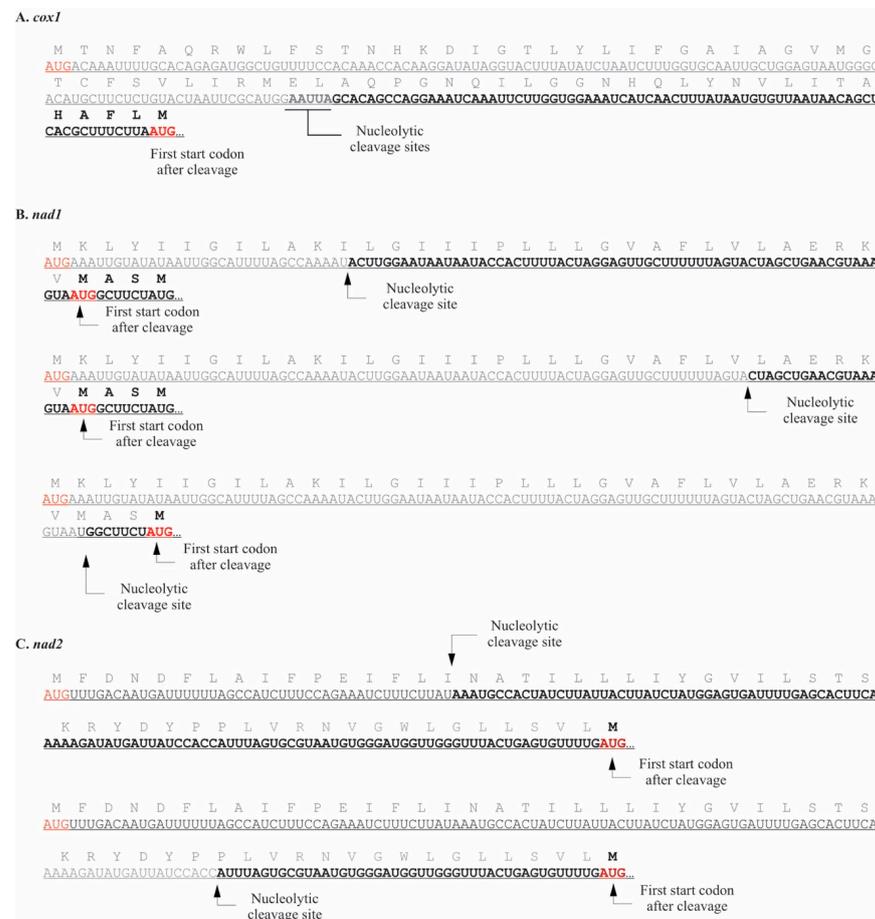
Figure 5. The 5′ termini and UTRs of three *C. vulgaris* mitochondrial mRNAs where the 5′ terminus occurred downstream of predicted start codons. Start codons are red and truncated portions of coding regions grey. (**A**). For *cox1* several 5′ termini were detected within a 5-nucleotide region. (**B**). Three 5′ termini were detected for *nad1*. (**C**). Two different termini were detected for *nad2*.

Table 4. Length of the most abundant 3′ UTRs found in *Chara vulgaris* and those with a non-template polynucleotide addition.

| Gene | Circular RT-PCR | | PacBio | | |
	Length of 3′ UTR from Stop Codon	UTRs with polyA Addition	Length of 3′ UTR from Stop Codon	Genes with an Adjacent tRNA (Distance from Stop Codon in nt)	Evidence of RNA Secondary Structure Immediately Downstream of 3′ Terminus
atp6	79 81 83	+ + +	nd	459	+
atp9	52–57 58	+ -	56	92	-
cob	54 68	- +	121		-
cox1	73	-	73		
cox2	51 82 107 162	- - - -	nd		-
cox3	23 51 56	+ - -	52	6056	-

Table 4. *Cont.*

	Circular RT-PCR		PacBio		
Gene	Length of 3′ UTR from Stop Codon	UTRs with polyA Addition	Length of 3′ UTR from Stop Codon	Genes with an Adjacent tRNA (Distance from Stop Codon in nt)	Evidence of RNA Secondary Structure Immediately Downstream of 3′ Terminus
nad1	20 67 70 141	- - + -	nd		+
nad2	87	+	nd	1189	+
nad4	10 13 21	+ - -	nd		-
nad4L	0 29	- +	nd	2508	-
nad5	nd	nd	nd		
nad6	nd	nd	nd	5	

nd: no data.

A. *atp6*
```
...UAACCACUUUCACUUCGUGGCCUCUGGUGACACAACCGCCGCGAAGCGUUCAUCUUUCCCACCCACCCCAGGGUGGGGUGGU-polyA
...UAACCACUUUCACUUCGUGGCCUCUGGUGACACAACCGCCGCGAAGCGUUCAUCUUUCCCACCCACCCCAGGGUGGGGUGGUGU-polyA
...UAACCACUUUCACUUCGUGGCCUCUGGUGACACAACCGCCGCGAAGCGUUCAUCUUUCCCACCCACCCCAGGGUGGGGUGGUGUCG-polyA
```

B. *atp9*
```
...UAAUGAAAUUAUGUACCCGUAAAAAAUGGCCGGGAAAUCCUGGCCUUACGGGUACCACCCU
...UAAUGAAAUUAUGUACCCGUAAAAAAUGGCCGGGAAAUCCUGGCCUUACGGGUACCACCC
```
polyA additions occurred within this region

C. *cob*
```
...UAGCUUUCAUUUUGAAUUGAAUGUCAUGCAUGAUAUGUACCAAUUACCCACCCGUUCGGG
...UAGCUUUCAUUUUGAAUUGAAUGUCAUGCAUGAUAUGUACCAAUUACCCACCCGUUCGGGUGGGGUUUCAA-polyA
```

D. *cox1*
```
...UAAUUGUUGUUAUAAGAAAGUAGUCCGUUUUUAUUAUAUUACAUUGCAUUGCCCCCAUUUAACUUGGGGGCAACGA
```

E. *cox2*
```
                             cox3
...UAAAAAACAAAGGAGCGAAGCAAUGAGUAUGAGUUCUCAAAAGCAUCCUUAUCA
...UAAAAAACAAAGGAGCGAAGCAAUGAGUAUGAGUUCUCAAAAGCAUCCUUAUCAUUUAGUUGAUCCAAGUCCAUGGCCUAUUUUU
...UAAAAAACAAAGGAGCGAAGCAAUGAGUAUGAGUUCUCAAAAGCAUCCUUAUCAUUUAGUUGAUCCAAGUCCAUGGCCUAUUUUUGGUUCA
CUGGGAGCUUUGGCAAGUA
...UAAAAAACAAAGGAGCGAAGCAAUGAGUAUGAGUUCUCAAAAGCAUCCUUAUCAUUUAGUUGAUCCAAGUCCAUGGCCUAUUUUUGGUUCA
CUGGGAGCUUUGGCAAGUACCAUUGGUGGUGUUAU
```

F. *cox3*
```
...UAACACACACUUACCCCGCCCCGGGG-polyA
...UAACACACACUUACCCCGCCCCGGGGCCCUUUACUUUAGGGCGCCGGGUCCACA
...UAACACACACUUACCCCGCCCCGGGGCCCUUUACUUUAGGGCGCCGGGUCCACAGAACA
```

G. *nad1*
```
...UAACUUCACUUUGUUAAGUAGUGUG
...UAACUUCACUUUGUUAAGUAGUGAUUCAUAUAGAAUAGAAAGGUGCCCACUAUAGACGUGGGGUGCCACUU
...UAACUUCACUUUGUUAAGUAGUGAUUCAUAUAGAAUAGAAAGGUGCCCACUAUAGACGUGGGGUGCCACUUCGC-polyA
...UAACUUCACUUUGUUAAGUAGUGAUUCAUAUAGAAUAGAAAGGUGCCCACUAUAGACGUGGGGUGCCACUUCGCCAGAUGUAUUGUAUCU
ACUUCUACUAGCUACGCUUAACAAAGAAAGAAGUUCCCACUUCCUUCCUUCCUUC
```

H. *nad2*
```
...UAAGUUUCGAGCGGAGACGAAACUUCGUAGUAAUAAAAAUAAAUAUAAAUACGCGCAGCUAACUAAAUGCCAUAAAGGCUGCGCGUACUCG-polyA
```

I. *nad4*
```
...UAAAGGGCCUUUG-polyA
...UAAAGGGCCUUUGUAU
...UAAAGGGCCUUUGUAUUUGACCGG
```

J. *nad4L*
```
...UAA
...UAAGAAAGGCGCCUAAAAAAAGAGGCGCCUUAG-polyA
```

Figure 6. The 3′ termini and UTRs of ten *C. vulgaris* mitochondrial mRNAs. Stop codons are blue. (**A**). Three termini were detected for *atp6*, a portion of each had a poly(A) addition. (**B**). Several termini were detected for *atp9*. The longer one (upper) had no oligonucleotide additions, while those lacking the terminal uracil had poly(A) additions within a 6 nt region. (**C**). Two *cob* termini were detected and the shorter one had no detectable oligonucleotide additions while the longer (lower) one did. (**D**). A single terminus was detected for *cob* with no oligonucleotide additions. (**E**). Four termini were detected for *cox2*. All had a portion of the adjacent *cox3* gene and no oligonucleotide additions. (**F**). Three *cox3* termini were detected. A portion of the shortest UTR had poly(A) tails, while the two longer ones did not. (**G**). Four termini were detected for *nad1* but only one had a poly(A) tail. (**H**). A single terminus was detected for *nad2* and a portion of them had a poly(A) tail. (**I**). Three termini were detected for *nad4* and a portion of the shortest had a poly(A) tail. (**J**). Two termini were detected for *nad4L*. One occurred directly adjacent to the stop codon. A portion of the longer UTR had a poly(A) addition.

Since tRNA placement was found to be important for 3′ maturation of some *P. duplex* genes, the distances of mature 3′ termini from the stop codons of genes with an adjacent tRNA were compared for *C. vulgaris* (Table 4). For the twelve genes used in this study, none had 3′ termini formed from the removal of an adjacent tRNA. The presence of RNA secondary structure (t-elements) immediately downstream of 3′ termini was also checked, and three genes (*atp6*, *nad1*, and *nad2*) had potential stem-loop structures adjacent to their mature 3′ termini.

In *C. vulgaris*, circularized full length coding regions were detected for five genes. Circularized *nad2* transcripts matched the 3′ and 5′ termini detected in earlier experiments (Figures 5C and 6H), suggesting that only naturally circularized transcripts were detected for this gene. For the genes *atp9*, *cox1*, *cox2*, and *nad4L*, circularized variants differed from the artificially circularized transcripts analyzed in previous experiments and are represented in Figure 7. For genes *cox2*, *cox3*, *nad1*, and *nad4*, only fragments of coding regions were found circularized. In *C. vulgaris*, there were no repeat motifs associated with the ligation sites and no polyA additions detected in the circularized transcripts.

Figure 7. Naturally circularized mRNAs found in *C. vulgaris* mitochondria. Start codons are colored red, stop codons blue, and truncated portions of coding regions grey. (**A**). Two circularized versions of full-length transcripts were detected for *atp9*. (**B**). Two circularized *cox1* transcripts were detected. Both had the same 5′ end but different termini on the 3′ ends. (**C**). A single circularized version of the *cox2a* transcript was detected. (**D**). Two circularized full-length *nad4L* transcripts of different lengths were detected.

3. Discussion

Green algae form two discrete clades, chlorophytes and charophytes [65]. Chlorophytes are morphologically diverse, cosmopolitan and contain the majority of green algae. Chlorophyte chondriomes are nearly as diverse as their morphologies, ranging from 15,500–66,000 bp with few introns and highly variable intergenic spaces, gene content, and gene synteny [60,66–75]. The knowledge of transcription and post-transcriptional processing events in chlorophytic mitochondria is limited to the model system *C. reinhardtii*, where mature mRNAs have no 5′ UTR and relatively short 3′ termini that may be poly-cytidylated. *C. reinhardtii* has a very small linear chondriome that is not representative of the majority of algae, and it is unknown if the post-transcriptional processing characteristics are hallmarks of its reduced chondriome or traits conserved among chlorophytes. Charophytes diverged from chlorophytes a billion or more years ago and have biochemical and morphological synapomorphies shared only with embryophytes [76]. The relatively few extant species belong to six monophyletic clades and are the closest living algal relatives of land plants [63,77]. The chondriomes of charophytes range widely from 56,500 bp to >201,000 bp, with highly variable gene order, density, and intron placement [63].

3.1. Mitochondrial Processing in P. duplex Resembles That Seen in C. reinhardtii

We mapped the 5′ and 3′ termini of presumably mature *P. duplex* mitochondrial mR-NAs and found the 5′ UTRs to be very short or non-existent, the 3′ UTRs of varying lengths, the polycytidylation of 3′ termini, and circularized full-length mRNAs. In the only other chlorophyte where mitochondrial RNA processing has been documented, *C. reinhardtii*, individual mRNAs are endonucleolytically cleaved from primary transcripts directly adjacent to the start codons, leaving no 5′ UTR. The remaining intergenic region becomes the 3′ UTR of an adjacent gene and is of varying lengths, presumably due to exonucleolytic processing [56]. It has also been demonstrated that non-template poly-cytidylation occurs on the 3′ termini of *C. reinhardtii*, and that this phenomenon may be limited to the Phylum Chlorophyta since it has been found in representative species from Peridinophyceae, Prasinophyceae, Trebouxiophyceae, and Chlorophyceae, but not in a red alga (*Chondrus crispus* Stackhouse), a glaucophyte (*Cyanophora paradoxa* Korshikov), or embryophytes (*Physcomitrella patens* (Hedw.) Bruch & Schimp., *Arabidopsis thaliana* (L.) Heynh., and *Solanum tuberosum* L.) [55]. *Chlamydomonas* mitochondrial mRNAs are also 3′ polyuridylylated and polyadenylated [54–57], but this was not observed in *P. duplex*.

Circular mRNAs are a common phenomenon across the biological spectrum, but their purpose has been difficult to determine [78]. In algal mitochondria mRNA circularization has been demonstrated in *Chlamydomonas*, where it appears to create translatable mRNAs and was hypothesized to be linked to polycytidylation [56]. In *P. duplex*, full-length coding regions were found to be circularized for seven of the twelve transcripts that were analyzed. The remaining five yielded no data so we are uncertain if they form circular transcripts. Circularization for five of the transcripts coincided with tandemly repeated template-derived motifs, AU repeats, GAACGAA, and GCGUCUU, which was not reported in *C. reinhardtii*. RNA circularization among nuclear genes occurs by way of an intron-exon back-splicing mechanism involving protein factors and conserved cis-elements [79,80], and perhaps a similar mechanism involving repeated elements occurs in mitochondria. The co-incidence of poly-cytidylation and circularization in *C. reinhardtii* led to a hypothesis that the poly(C) acted as a cis-element for circularization. In *P. duplex*, only one transcript, *atp9*, had circular mRNAs with full-length coding regions and a poly(C) addition. The other two transcripts with poly(C) additions, *atp6* and *cob*, were fragments of the coding region, suggesting that the poly(C) addition is not universally linked to the creation of translatable circular mRNAs and may not act as a signal for circularization.

These data demonstrate that mitochondrial mRNA processing is conserved in the Chlorophyceae algal clade. The polycytidylation of mitochondrial mRNAs has already been shown to be conserved across the chlorophytic algae [55]. This study extends the similarities of mitochondrial mRNA processing to include the absence of a 5′ UTR, despite *P. duplex* having larger intergenic regions than *C. reinhardtii*. Our data suggest that the cleavage of tRNAs plays a pivotal role in the maturation of two mRNAs (*cob* and *atp9*) in *P. duplex*, which is consistent with mitochondrial transcripts in all systems studied to date, from *C. reinhardtii* to humans. The difference is the size of the chondriome and the lengths of the intergenic regions. We found that the removal of tRNAs in *P. duplex* creates mature 3′ ends that, at least in two cases, require no further processing other than polycytidylation. The 5′ ends, however, are further processed until no 5′ UTR remains (Figure 4). The lack of a 5′ UTR and the accompanying translation mechanism required is not only due to the compressed chondriome of *C. reinhardtii*, but also its purposeful removal prior to translation. *Pediastrum* is in the order Sphaeropleales and is a sister clade to Volvocales, which contains *Chlamydomonas* [81], so it is possible that these processing events could be limited to those two Orders; however, we hypothesize that these processes are conserved across the Chlorophyceae.

3.2. Mitochondrial Processing in C. vulgaris Resembles That Seen in Embryophytes.

We also analyzed the mRNA termini of the mitochondrial mRNAs of *C. vulgaris* and found the 5′ and 3′ UTRs to be much longer and more variable than those observed in

P. duplex and *C. reinhardtii*. We also detected non-template polyadenylation of 3′ termini, the possibility of RNA secondary structures in processed 3′ UTRs, and the possibility of t-elements downstream of some 3′ termini. The *C. vulgaris* chondriome is one of the smaller among charophytic algae, 67,737 bp, with intergenic region sizes more similar to Chlorophyeae than higher plants, yet with gene content, gene synteny, and intron placement very similar to the much larger bryophyte chondriomes [62,63]. After post-transcriptional processing, we found that the 3′ and 5′ UTR lengths in *C. vulgaris* were more similar to those found in embryophytes than chlorophytes. In embryophytes, 5′ termini are of varying lengths, ranging from dozens to thousands of nucleotides in length, presumably due to a stepwise 5′ maturation process [29–36,82–86]. Among embryophytes, 3′ UTR lengths tend to be more consistent in length and mature mRNAs lack non-template oligonucleotide additions other than polyadenylation, which marks the mRNA for degradation [44–47]. Embryophyte mitochondria utilize tRNAs and RNA secondary structure (t-elements) to guide endonucleolytic cleavage [31–33]. The presence of relatively long 5′ and 3′ UTRs, the polyadenylation of some 3′ termini, the possibility that secondary structures may guide the formation of some 3′ termini, and the possibility that secondary structures may form in the 3′ UTRs of *C. vulgaris* mitochondrial mRNAs to presumably provide a target for 3′ specific RPFs demonstrates that their RNA processing is more similar to embryophytes than chlorophytes. One recent proteomic survey of *Arabidopsis* mitochondria estimated that 14.9% of the nuclear encoded mitochondrial-targeted proteome is devoted to mRNA processing [87]. Our results suggest that the RNA processing mechanisms and perhaps the large numbers of nuclear-encoded RPFs that locate to the mitochondria in higher plants were present in the common ancestor of all streptophytes.

In *C. vulgaris*, circularized full length coding regions were detected for five genes and only fragments were found for four other genes. Unlike *P. duplex*, there were no repeat motifs associated with the ligation sites and no polyA additions detected in the circularized transcripts. Among embryophyte mitochondria, circularized RNAs were characterized in *Hordeum vulgare* and *Arabidopsis*, but all of the examples were fragmented coding regions [88,89]. Our data suggest that the circularization of mRNAs in *C. vulgaris* may be a hybrid of chlorophytes and embryophytes.

4. Materials and Methods

4.1. Cultures

The *P. duplex* strain used in this study was morphologically identified and its chondriome sequenced and archived in GenBank (MK895949), as described in [58]. It has been maintained in Bold's basal medium [90] in an Erlenmeyer flask on a lab bench with ambient light (adjacent to a window). The *C. vulgaris* strain was collected from a pond on the University of Virginia's College at Wise campus in 2015 and cultured in an aquarium containing mud substrate collected from the same pond. This culture has been maintained in a greenhouse and supplemented with commercial plant fertilizer [59]. Its chondriome has been fully sequenced and is nearly identical to the one archived by Turmel et al. [62] (NC-005255).

4.2. RNA Extraction and CircRT-PCR

To promote RNA production, *P. duplex* cultures were incubated in a shaking incubator (50 rpm) at constant temperature (28 °C) and artificial light (500 μE m^{-2} s^{-1}) for several hours prior to extraction. Cells were pelleted using a Beckman-Coulter Avanti JXN-30 centrifuge, and pellets were resuspended in Qiagen's RNeasy extraction buffer (Germantown, MD, USA) and transferred to a bead beater tube. Cells were lysed by bead beating on a vortexer for five minutes. The lysate was then taken through the remaining Qiagen RNeasy protocol with the optional DNase step. *C. vulgaris* tissues were flash frozen in liquid nitrogen and ground with a mortar and pestle followed by RNA extraction using a Qiagen RNeasy kit, as described in Cahoon et al. [59]. All RNA samples were quantified using a NanoDrop Lite (Thermo-Fisher, Waltham, MA, USA) and stored at −80 °C.

Primers for cDNA synthesis and PCR (Supplemental Table S1) were designed using Primer3 (https://primer3.org) and synthesized by Integrated DNA Technologies (Coralville, IA, USA). RT-PCR was completed as described in Meade et al. [91]. Briefly, RNAs were artificially circularized using 2 µg of total RNA and T4 RNA ligase (New England Biolabs, Ipswich, MA, USA). cDNAs were synthesized from the circularized RNAs using the R1 primers and MMLV Reverse Transcriptase (Promega, Madison, WI, USA). These cDNAs were then used as template, along with the R1 and L1 primers to PCR amplify the 3′–5′ junctions of each transcript using Phusion DNA polymerase (ThermoFisher, Waltham, MA, USA). The products of these PCR reactions were diluted 10-fold and used as template for a second round of PCR using primers R2 and L2. This process was completed twice with RNA extracted at two different times to represent independent replicates. Naturally circularized mRNAs were detected producing cDNA directly from 2 µg of total RNA without T4 ligase treatment followed by two rounds of PCR as described above. The naturally circularized mRNA process was also completed twice.

4.3. Sequencing and Analysis

The independent replicates of the secondary PCR products for both *P. duplex* and *C. vulgaris* were deep sequenced, separately, using Genewiz's Amplicon EZ Illumina MiSeq service (South Plainfield, NJ, USA). Sequences were analyzed using Geneious Prime software (Biomatters, Auckland, New Zealand). Initially, sequences were matched to each gene using the Map-to-Reference function. These sequences were visually inspected, for 3′ poly-nucleotide additions, and sequence motifs. The 3′–5′ junction sites were identified using NCBI's BLAST align two or more sequences feature (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

For PacBio Iso-Seq, ~1 g of total RNA from each species was polyadenylated using Lucigen's (Middleton, WI, USA) Poly(A) Polymerase Tailing kit according to the manufacturer's protocol. The tailed samples were cleaned using Qiagen's RNeasy kit. Samples were shipped to GeneWiz on dry ice for PacBio library preparation and sequencing. This process was completed once. Reads were aligned to the *P. duplex* and *C. vulgaris* chondriomes with Geneious Prime software using the Map-to-Reference function.

Logo plots were generated using the web-based service https://weblogo.berkeley.edu/logo.cgi. RNA secondary structures were predicted using RNAfold (http://rna.tbi.univie.ac.at//cgi-bin/RNAWebSuite/RNAfold.cgi).

5. Conclusions

We present evidence of mitochondrial mRNA processing from two green algae from the chlorophyte and streptophyte lineages with similarly sized circular mitochondrial genomes. The primary differences were the absence of secondary structures (t-elements) in the *P. duplex* chondriome that are important for processing in streptophytes, and the removal of the 5′ UTR in chlorophytes but not streptophytes. We hypothesize that t-elements were gained and the 5′ UTR processing lost in the common ancestor of the streptophytic algae. We also confirm the polycytidylation of the 3′ termini of *P. duplex* but not *C. vulgaris*, which is consistent with the theory that poly(C) addition is limited to chlorophytes.

Author Contributions: Conceptualization, A.B.C., K.M.M., G.C.R.P., M.J.M.; methodology, A.B.C., K.M.M., G.C.R.P., M.J.M.; investigation, G.C.R.P., M.J.M.; resources, A.B.C., K.M.M.; writing—original draft preparation, A.B.C., G.C.R.P., M.J.M.; writing—review and editing, A.B.C., K.M.M., G.C.R.P., M.J.M.; visualization, A.B.C., G.C.R.P.; supervision, A.B.C., K.M.M.; project administration, A.B.C. All authors have read and agreed to the published version of the manuscript.

References

1. Rose, R.J.; Tiew, T.W.-Y.; Dashek, W. The mitochondrion. In *Plant Cells and Their Organelles*; Dashek, W.V., Miglani, G.S., Eds.; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2017.
2. Archibald, J.M. Endosymbiosis and eukaryotic cell evolution. *Curr. Biol.* **2015**, *25*, R911–R921. [CrossRef]
3. López-García, P.; Eme, L.; Moreira, D. Symbiosis in eukaryotic evolution. *J. Theor. Biol.* **2017**, *434*, 20–33. [CrossRef] [PubMed]
4. Wang, Z.; Wu, M. An integrated phylogenomic approach toward pinpointing the origin of mitochondria. *Sci. Rep.* **2015**, *5*, 7949. [CrossRef] [PubMed]
5. Johnston, I.G.; Williams, B.P. Evolutionary inference across eukaryotes identifies specific pressures favoring mitochondrial gene retention. *Cell Syst.* **2016**, *2*, 101–111. [CrossRef] [PubMed]
6. D'Souza, A.R.; Minczuk, M. Mitochondrial transcription and translation: Overview. *Essays Biochem.* **2018**, *62*, 309–320. [CrossRef]
7. Tiranti, V.; Savoia, A.; Forti, F.; D'Apolito, M.F.; Centra, M.; Rocchi, M.; Zeviani, M. Identification of the gene encoding the human mitochondrial RNApolymerase (h-mtRPOL) by cyberscreening of the expressed sequence tags database. *Hum. Mol. Genet.* **1997**, *6*, 615–625. [CrossRef]
8. Falkenberg, M.; Gaspari, M.; Rantanen, A.; Trifunovic, A.; Larsson, N.G.; Gustafsson, C.M. Mitochondrial transcription factors B1 and B2 activate transcription of human mtDNA. *Nat. Genet.* **2002**, *31*, 289–294. [CrossRef]
9. Kanki, T.; Nakayama, H.; Sasaki, N.; Takio, K.; Alam, T.I.; Hamasaki, N.; Kang, D. Mitochondrial nucleoid and transcription factor A. *Ann. N. Y. Acad. Sci.* **2004**, *1011*, 61–68. [CrossRef]
10. Montoya, J.; Christianson, T.; Levens, D.; Rabinowitz, M.; Attardi, G. Identification of initiation sites for heavy-strand and light-strand transcription in human mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 7195–7199. [CrossRef]
11. Montoya, J.; Gaines, G.L.; Attardi, G. The pattern of transcription of the human mitochondrial rRNA genes reveals two overlapping transcription units. *Cell* **1983**, *34*, 151–159. [CrossRef]
12. Chang, D.D.; Clayton, D.A. Precise identification of individual promoters for transcription of each strand of human mitochondrial DNA. *Cell* **1984**, *36*, 635–643. [CrossRef]
13. Gao, S.; Tian, X.; Chang, H.; Sun, Y.; Wu, Z.; Cheng, Z.; Dong, P.; Zhao, Q.; Ruan, J.; Bu, W. Two novel lncRNAs discovered in human mitochondrial DNA using PacBio full-length transcriptome data. *Mitochondria* **2018**, *38*, 41–47. [CrossRef] [PubMed]
14. Minczuk, M.; He, J.; Duch, A.M.; Ettema, T.J.; Chlebowski, A.; Dzionek, K.; Nijtmans, L.G.J.; Huynen, M.A.; Holt, I.J. TEFM (c17orf42) is necessary for transcription of human mtDNA. *Nucleic Acids Res.* **2011**, *39*, 4284–4299. [CrossRef] [PubMed]
15. Ojala, D.; Montoya, J.; Attardi, G. tRNA punctuation model of RNA processing inhuman mitochondria. *Nature* **1981**, *290*, 470–474. [CrossRef]
16. Montoya, J.; Ojala, D.; Attardi, G. Distinctive features of the 5′-terminal sequences of the human mitochondrial mRNAs. *Nature* **1981**, *290*, 465–470. [CrossRef]
17. Temperley, R.J.; Wydro, M.; Lightowlers, R.N.; Chrzanowska-Lightowlers, Z.M. Human mitochondrial mRNAs—Like members of all families, similar but different. *Biochim. Biophys. Acta* **2010**, *1797*, 1081–1085. [CrossRef]
18. Kummer, E.; Leibundgut, M.; Rackham, O.; Lee, R.G.; Boehringer, D.; Filipovska, A.; Ban, N. Unique features of mammalian mitochondrial translation initiation revealed by cryo-EM. *Nature* **2018**, *560*, 263–267. [CrossRef] [PubMed]
19. Nagaike, T.; Suzuki, T.; Katoh, T.; Ueda, T. Human mitochondrial mRNAs are stabilized with polyaenylation regulated by mitochondria-specific poly(A) polymerase and polynucleotide phosphorylase. *J. Biol. Chem.* **2005**, *280*, 19721–19727. [CrossRef]
20. Slomovic, S.; Lauer, D.; Geiger, D.; Schuster, G. Polyadenylation and degradation of human mitochondrial RNA: The prokaryotic past leaves its mark. *Mol. Cell. Biol.* **2005**, *25*, 6427–6435. [CrossRef] [PubMed]
21. Kuznetsova, I.; Siira, S.J.; Shearwood, A.-M.J.; Ermer, J.A.; Filipovska, A.; Rackham, O. Simultaneous processing and degradation of mitochondrial RNAs revealed by circularized RNA sequencing. *Nucleic Acids Res.* **2017**, *45*, 5487–5500. [CrossRef] [PubMed]
22. Mance, L.G.; Mawla, I.; Shell, S.M.; Cahoon, A.B. Mitochondrial mRNA fragments are circularized in a human HEK cell line. *Mitochondrion* **2020**, *51*, 1–6. [CrossRef]
23. Gualberto, J.M.; Newton, K.J. Plant mitochondrial genomes: Dynamics and mechanisms of mutation. *Annu. Rev. Plant Biol.* **2017**, *68*, 225–252. [CrossRef] [PubMed]
24. Hedtke, B.; Börner, T.; Weihe, A. One RNA polymerase serving two genomes. *EMBO Rep.* **2000**, *1*, 435–440. [CrossRef]
25. Kühn, K.; Richter, U.; Meyer, E.H.; Delannoy, E.; de Longevialle, A.F.; O'Toole, N.; Börner, T.; Millar, A.H.; Small, I.D.; Whelan, J. Phage-type RNA polymerase RPOTmp performs gene-specific transcription in mitochondria of *Arabidopsis thaliana*. *Plant Cell* **2009**, *21*, 2762–2779. [CrossRef]
26. Roberti, M.; Polosa, P.L.; Bruni, F.; Manzari, C.; Deceglie, S.; Gadaleta, M.N.; Cantatore, P. The MTERF family proteins: Mitochondrial transcription regulators and beyond. *Biochim. Biophys. Acta* **2009**, *1787*, 303–311. [CrossRef] [PubMed]
27. Kühn, K.; Weihe, A.; Börner, T. Multiple promoters are a common feature of mitochondrial genes in *Arabidopsis*. *Nucleic Acids Res.* **2005**, *33*, 337–346. [CrossRef] [PubMed]

28. Perrin, R.; Meyer, E.H.; Zaepfel, M.; Kim, Y.J.; Mache, R.; Grienenberger, J.M.; Gualberto, J.M.; Gagliardi, D. Two exoribonucleases act sequentially to process mature 3′-ends of atp9 mRNAs in *Arabidopsis* mitochondria. *J. Biol. Chem.* **2004**, *279*, 25440–25446. [CrossRef] [PubMed]

29. Zhang, Y.; Huang, X.; Zou, J.; Liu, Y.; Lian, T.; Nian, H. Major contribution of transcription initiation to 5′-end formation of mitochondrial steady-state transcripts in maize. *RNA Biol.* **2019**, *16*, 104–117. [CrossRef]

30. Binder, S.; Stoll, K.; Stoll, B. Maturation of 5′ ends of plant mitochondrial RNAs. *Physiol. Plant* **2016**, *157*, 280–288. [CrossRef] [PubMed]

31. Forner, J.; Weber, B.; Thuss, S.; Wildum, S.; Binder, S. Mapping of mitochondrial mRNA termini in *Arabidopsis thaliana*: T-elements contribute to 5′ and 3′ end formation. *Nucleic Acids Res.* **2007**, *35*, 3676–3692. [CrossRef] [PubMed]

32. Hanic-Joyce, P.J.; Spencer, D.F.; Gray, M.W. In vitro processing of transcripts containing novel tRNA-like sequences ('t-elements') encoded by wheat mitochondrial DNA. *Plant Mol. Biol.* **1990**, *15*, 551–559. [CrossRef] [PubMed]

33. Dombrowski, S.; Brennicke, A.; Binder, S. 3′-inverted repeats in plant mitochondrial mRNAs are processing signals rather than transcription terminators. *EMBO J.* **1997**, *16*, 5069–5076. [CrossRef]

34. Gobert, A.; Gutmann, B.; Taschner, A.; GöBringer, M.; Holzmann, J.; Hartmann, R.K.; Rossmanith, W.; Giegé, P. A single *Arabidopsis* organellar protein has RNase P activity. *Nat. Struct. Mol. Biol.* **2010**, *17*, 740–744. [CrossRef] [PubMed]

35. Gutmann, B.; Gobert, A.; Giegé, P. PRORP proteins support RNase P activity in both organelles and the nucleus in *Arabidopsis*. *Genes Dev.* **2012**, *26*, 1022–1027. [CrossRef] [PubMed]

36. Stoll, B.; Binder, S. Two NYN domain containing putative nucleases are involved in transcript maturation in *Arabidopsis* mitochondria. *Plant J.* **2016**, *85*, 278–288. [CrossRef] [PubMed]

37. Jonietz, C.; Forner, J.; Hölzle, A.; Thuss, S.; Binder, S. RNA PROCESSING FACTOR2 is required for 5′ end processing of nad9 and cox3 mRNAs in mitochondria of *Arabidopsis thaliana*. *Plant Cell* **2010**, *22*, 443–453. [CrossRef]

38. Hölzle, A.; Jonietz, C.; Törjek, O.; Altmann, T.; Binder, S.; Forner, J. A RESTORER OF FERTILITY-like PPR gene is required for 5′-end processing of the nad4 mRNA in mitochondria of *Arabidopsis thaliana*. *Plant J.* **2011**, *65*, 65,737–744. [CrossRef]

39. Hauler, A.; Jonietz, C.; Stoll, B.; Stoll, K.; Braun, H.-P.; Binder, S. RNA PROCESSING FACTOR 5 is required for efficient 5′ cleavage at a processing site conserved in RNAs of three different mitochondrial genes in *Arabidopsis thaliana*. *Plant J.* **2013**, *74*, 593–604. [CrossRef]

40. Haili, N.; Arnal, N.; Quadrado, M.; Amiar, S.; Tcherkez, G.; Dahan, J.; Briozzo, P.; Colas des Francs-Small, C.; Vrielynck, N.; Mireau, H. The pentatricopeptide repeat MTSF1 protein stabilizes the nad4 mRNA in *Arabidopsis* mitochondria. *Nucleic Acids Res.* **2013**, *41*, 6650–6663. [CrossRef]

41. Hammani, K.; Giegé, P. RNA metabolism in plant mitochondria. *Trends Plant Sci.* **2014**, *19*, 380–389. [CrossRef]

42. Brown, G.G.; des Francs-Small, C.C.; Ostersetzer-Biran, O. Group II intron splicing factors in plant mitochondria. *Front. Plant Sci.* **2014**, *5*, 35. [CrossRef] [PubMed]

43. Small, I.D.; Schallenberg-Rüdinger, M.; Takenaka, M.; Mireau, H.; Ostersetzer-Biran, O. Plant organellar RNA editing: What 30 years of research has revealed. *Plant J.* **2020**, *101*, 1040–1056. [CrossRef] [PubMed]

44. Gagliardi, D.; Leaver, C.J. Polyadenylation accelerates the degradation of the mitochondrial mRNA associated with cytoplasmic male sterility in sunflower. *EMBO J.* **1999**, *18*, 3757–3766. [CrossRef] [PubMed]

45. Lupold, D.S.; Caoile, A.G.; Stern, D.B. Polyadenylation occurs at multiple sites in maize mitochondrial cox2 mRNA and is independent of editing status. *Plant Cell* **1999**, *11*, 1565–1578. [CrossRef] [PubMed]

46. Kühn, J.; Tengler, U.; Binder, S. Transcript lifetime is balanced between stabilizing stem-loop structures and degradation-promoting polyadenylation in plant mitochondria. *Mol. Cell. Biol.* **2001**, *21*, 731–742. [CrossRef] [PubMed]

47. Schuster, G.; Stern, D. RNA polyadenylation and decay in mitochondria and chloroplasts. *Prog. Mol. Biol. Transl. Sci.* **2009**, *85*, 393–422. [CrossRef]

48. Harris, E.H. *The Chlamydomonas Sourcebook*, 2nd ed.; Harris, E.H., Stern, D.B., Witman, G.B., Eds.; Academic Press: Oxford, UK, 2009.

49. Gray, M.W.; Boer, P.H. Organization and expression of algal (*Chlamydomonas reinhardtii*) mitochondrial DNA. *Phil. Trans. R. Soc. Lon. B Biol. Sci.* **1988**, *319*, 135–147. [CrossRef]

50. Laflamme, M.; Lee, R.W. Mitochondrial genome conformation among CW-group chlorophycean algae. *J. Phycol.* **2003**, *39*, 213–220. [CrossRef]

51. Mallet, M.A.; Lee, R.W. Identification of three distinct *Polytomella* lineages based on mitochondrial DNA features. *J. Eukaryot. Microbiol.* **2006**, *53*, 79–84. [CrossRef]

52. Tracy, R.L.; Stern, D.B. Mitochondrial transcription initiation: Promoter structures and RNA poymerases. *Curr. Genet.* **1995**, *28*, 205–216. [CrossRef]

53. Duby, F.; Cardol, P.; Matgne, R.F.; Remacle, C. Structure of the telomeric ends of mt DNA, transcriptional analysis and complex I assembly in the dum24 mitochondrial mutant of *Chlamydomonas reinhardtii*. *Mol. Genet. Genom.* **2001**, *266*, 109–114. [CrossRef]

54. Zimmer, S.L.; Schein, A.; Zipor, G.; Stern, D.B.; Schuster, G. Polyadenylation in *Arabidopsis* and *Chlamydomonas* organelles: The input of nucleotidyltransferases, poly(A) polymerases and polynucleotide phosphorylase. *Plant J.* **2009**, *59*, 88–99. [CrossRef] [PubMed]

55. Salinas-Giegé, T.; Cavaiuolol, M.; Cognat, V.; Ubrig, E.; Remacle, C.; Duchêne, A.-M.; Vallon, O.; Maréchal-Drouard, L. Poly-cytidylation of mitochondrial mRNAs in *Chlamydomonas reinhardtii*. *Nucleic Acids Res.* **2017**, *45*, 12963–12973. [CrossRef]

56. Cahoon, A.B.; Qureshi, A.A. Leaderless mRNAs are circularized in *Chlamydomonas reinhardtii* mitochondria. *Curr. Genet.* **2018**, *64*, 1321–1333. [CrossRef]

57. Gallaher, S.D.; Fitz-Gibbon, S.T.; Strenkert, D.; Purvine, S.O.; Pellegrini, M.; Merchant, S.S. High-throughput sequencing of the chloroplast and mitochondrion of *Chlamydomonas reinhardtii* to generate improved de novo assemblies, analyze expression patterns and transcript speciation, and evaluate diversity among laboratory strains and wild isolates. *Plant J.* **2017**, *93*, 545–565. [CrossRef]

58. Levy, S.; Schuster, G. Polyadenylation and degradation of RNA in the mitochondria. *Biochem. Soc. Trans.* **2016**, *44*, 1475–1482. [CrossRef]

59. Cahoon, A.B.; Nauss, J.A.; Stanley, C.D.; Qureshi, A. Deep Transcriptome Sequencing of Two Green Algae, *Chara vulgaris* and *Chlamydomonas reinhardtii*, Provides no Evidence of Organellar RNA Editing. *Genes* **2017**, *8*, 80. [CrossRef] [PubMed]

60. Farwagi, A.; Fučíková, K.; McManus, H.A. Phylogenetic patterns of gene rearrangements in four mitochondrial genomes from the green algal family Hydrodictyaceae (Sphaeropleales, Chlorophyceae). *BMC Genom.* **2015**, *16*, 826. [CrossRef] [PubMed]

61. Proulex, G.C.R.; Lor, B.; Manoylov, K.M.; Cahoon, A.B. The chloroplast and mitochondrial genomes of the green alga *Pediastrum duplex* isolated from central Georgia (USA). *Mitochondrial DNA Part B Resour.* **2019**, *4*, 3070–3071. [CrossRef]

62. Turmel, M.; Otis, C.; Lemieux, C. The mitochondrial genome of *Chara vulgaris*: Insights into the mitochondrial DNA architecture or the last common ancestor of green algae and land plants. *Plant Cell* **2003**, *15*, 1888–1903. [CrossRef] [PubMed]

63. Turmel, M.; Otis, C.; Lemieux, C. Tracing the evolution of streptophyte algae and their mitochondrial genome. *Genome Biol. Evol.* **2013**, *5*, 1817–1835. [CrossRef] [PubMed]

64. Greiner, S.; Lehwark, P.; Bock, R. OrganellarGenomeDraw (OGDRAW) version 1.3.1: Expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res.* **2019**, *47*, W59–W64. [CrossRef]

65. Leliaert, F.; Smith, D.R.; Moreau, H.; Herron, M.D.; Verbruggen, H.; Delwiche, C.F.; De Clerck, O. Phylogeny and molecular evolution of the green algae. *Crit. Rev. Plant Sci.* **2012**, *31*, 1–46. [CrossRef]

66. Boer, P.H.; Bonen, L.; Lee, R.W.; Gray, M.W. Genes for respiratory chain proteins and ribosomal RNAs are present on a 16-kilobase-pair DNA species from *Chlamydomonas reinhardtii* mitochondria. *Proc. Natl. Acad. Sci. USA* **1985**, 3340–3344. [CrossRef] [PubMed]

67. Turmel, M.; Lemieux, C.; Burger, G.; Lang, B.F.; Otis, C.; Plante, I.; Gray, M.W. The complete mitochondrial DNA sequences of *Nephroselmis olivacea* and *Pedinomonas minor*: Two radically different evolutoionary patterns within green algae. *Plant Cell* **1999**, *11*, 1717–1729. [CrossRef] [PubMed]

68. Nedelcu, A.M.; Lee, R.W.; Lemieux, C.; Gray, M.W.; Burger, G. The complete mitochondrial DNA sequence of *Scenedesmus obliquus* reflects an intermediate stage in the evolution of the green algal mitochondrial genome. *Genome Res.* **2000**, *10*, 819–831. [CrossRef] [PubMed]

69. Robbens, S.; Derelle, E.; Ferraz, C.; Wuyts, J.; Moreau, H.; Van de Peer, Y. The complete chloroplast and mitochondrial DNA sequences of *Ostreococcus tauri*: Organelle genomes of the smallest eukaryote are examples of compaction. *Mol. Biol. Evol.* **2007**, *24*, 956–968. [CrossRef] [PubMed]

70. Brouard, J.-S.; Otis, C.; Lemieux, C.; Turmel, M. Chloroplast DNA sequence of the green alga *Oedogonium cardiacum* (Chlorophyceae): Unique genome architecture, derived characters shared with the Chaetophorales and novel genes acquired through horizontal transfer. *BMC Genom.* **2008**, *9*, 290. [CrossRef]

71. Pombert, J.-F.; Otis, C.; Lemieux, C.; Turmel, M. The complete mitochondrial DNA sequence of the green alga *Pseudenclonium akinetum* (Ulvophyceae) highlights distinctive evolutionary trends in the Chlorophyta and suggests a sister-group relationship between the Ulvophyceae and Chlorophyceae. *Mol. Biol. Evol.* **2004**, *21*, 922–935. [CrossRef]

72. Pombert, J.-F.; Otis, C.; Turmel, M.; Lemieux, C. The mitochondrial genome of the prasinophyte *Prasinoderma colonial* reveals two *trans*-spliced group I introns in the large subunit rRNA gene. *PLoS ONE* **2013**, *8*, e84325. [CrossRef]

73. Smith, D.R.; Hamaji, T.; Olson, B.J.S.C.; Durand, P.M.; Ferris, P.; Michod, R.E.; Featherston, J.; Nozaki, H.; Keeling, P.J. Organelle genome complexity scales positively with organism size in volvocine green algae. *Mol. Biol. Evol.* **2013**, *30*, 793–797. [CrossRef] [PubMed]

74. Jeong, H.; Lim, J.-M.; Park, J.; Sim, Y.M.; Choi, H.-G.; Lee, J.; Jeong, W.-J. Plastid and mitochondrion genomic sequences from arctic *Chlorella* sp. ArM0029B. *BMC Genom.* **2014**, *15*, 286. [CrossRef] [PubMed]

75. Fučíková, K.; Lewis, P.O.; González-Halphen, D.; Lewis, L.A. Gene arrangement convergence, diverse intron content, and genetic code modifications in mitochondrial genomes of Sphaeropleales (Chlorophyta). *Genome Biol. Evol.* **2014**, *6*, 2170–2180. [CrossRef]

76. Nishiyama, T.; Sakayama, H.; de Vries, J.; Buschmann, H.; Saint-Marcoux, D.; Ullrich, K.K.; Haas, F.B.; Vanderstraeten, L.; Becker, D.; Lang, D.; et al. The *Chara* genome: Secondary complexity and implications for plant terrestrialization. *Cell* **2018**, *174*, 448–464. [CrossRef]

77. Karol, K.G.; McCourt, R.M.; Cimino, M.T.; Delwiche, C.F. The closest living relatives of land plants. *Science* **2001**, *294*, 2351–2353. [CrossRef] [PubMed]

78. Guria, A.; Sharma, P.; Natesan, S.; Pandi, G. Circular RNAs—The road less traveled. *Front. Mol. Biosci.* **2020**, *6*, 146. [CrossRef]

79. Starke, S.; Jost, I.; Rossbach, O.; Schneider, T.; Schriner, S.; Hung, L.-H.; Bindereif, A. Exon circularization requires canonical splice signals. *Cell Rep.* **2015**, *10*, 103–111. [CrossRef] [PubMed]

80. Bolha, L.; Ravnik-Glavač, M.; Glavač, D. Circular RNAs: Biogenesis, Function and a Role as Possible Cancer Biomarkers. *Int. J. Genom.* **2017**, *2017*, 6218353. [CrossRef]

81. Tippery, N.P.; Fučiková, K.; Lewis, P.O.; Lewis, L.A. Probing the monophyly of the Sphaeropleales (Chlorophyceae) using data from five genes. *J. Phycol.* **2012**, *48*, 1482–1493. [CrossRef]

82. Arnal, N.; Quadrado, M.; Simon, M.; Mireau, H. A restorer-of-fertility like pentatricopeptide repeat gene directs ribonucleolytic processing within the coding sequence of rps3-rpl16 and orf240a mitochondrial transcripts in *Arabidopsis thaliana*. *Plant J.* **2014**, *78*, 134–145. [CrossRef]

83. Jonietz, C.; Forner, J.; Hildebrandt, T.; Binder, S. RNA PROCESSING FACTOR 3 is crucial for the accumulation of mature ccmC transcripts in mitochondria of *Arabidopsis thaliana* accession Columbia. *Plant Physiol.* **2011**, *157*, 1430–1439. [CrossRef] [PubMed]

84. Stoll, B.; Stoll, K.; Steinhilber, J.; Jonietz, C.; Binder, S. Mitochondrial transcript length polymorphisms are a widespread phenomenon in *Arabidopsis thaliana*. *Plant Mol. Biol.* **2012**, *1*, 221–233. [CrossRef] [PubMed]

85. Stoll, B.; Zendler, D.; Binder, S. RNA processing factor 7 and polynucleotide phosphorylase are necessary for processing and stability of nad2 mRNA in *Arabidopsis* mitochondria. *RNA Biol.* **2014**, *11*, 968–976. [CrossRef] [PubMed]

86. Stoll, K.; Jonietz, C.; Binder, S. In *Arabidopsis thaliana* two co-adapted cyto-nuclear systems correlate with distinct ccmC transcript sizes. *Plant J.* **2015**, *81*, 247–257. [CrossRef]

87. Fuchs, P.; Rugen, N.; Carrie, C.; Elsässer, M.; Finkemeier, I.; Giese, J.; Hildebrandt, T.M.; Kühn, K.; Maurino, V.G.; Ruberti, C.; et al. Single organelle function and organization as estimated from *Arabidopsis* mitochondrial proteomics. *Plant J.* **2019**, *101*, 420–441. [CrossRef]

88. Darbani, B.; Noeparvar, S.; Borg, S. Identification of circular RNAs from parental genes involved in multiple aspects of cellular metabolism in barley. *Front. Plant Sci.* **2016**, *7*, 776. [CrossRef]

89. Sun, X.; Wang, X.; Ding, J.; Wang, Y.; Wang, J.; Zhang, X.; Che, Y.; Liu, Z.; Zhang, X.; Ye, J.; et al. Integrative analysis of *Arabidopsis thaliana* transcriptomics reveals intuitive splicing mechanism for circular RNA. *FEBS Lett.* **2016**, *590*, 3510–3516. [CrossRef]

90. Bischoff, H.W.; Bold, H.C. Phycological studies. IV. Some soil algae from Enchanted Rock and related algal species. *Univ. Tex. Publ.* **1963**, *6318*, 1–95.

91. Meade, M.J.; Proulex, G.C.R.; Manoylov, K.M.; Cahoon, A.B. Chloroplast MRNAs are 3′ Polyuridylylated in The Green Alga *Pithophora roettleri* (Cladophorales). *J. Phycol.* **2020**, *56*, 1124–1134. [CrossRef]