Does mitochondrial DNA replication in Chlamydomonas require a reverse transcriptase?

Introduction

Among the first green algal mitochondrial DNAs (mtDNAs) to be described in great detail was that of Chlamydomonas reinhardtii (Grant & Chiang, 1980; Gray & Boer, 1988; Vahrenholz et al., 1993). It is a 16 kb linear molecule containing fragmented rRNA genes and telomeres, arranged in an inverted-repeat orientation with 3’ single-stranded overhangs (Fig. 1a). Another intriguing feature of this genome is a free-standing open reading frame (1119 nt) encoding a reverse transcriptase-like protein (the rtl gene), which is believed to have originated from an ancient group II intron (Boer & Gray, 1988) but is no longer associated with one. Comparative genomics and transcriptomics have shown that rtl is under purifying selection (albeit weaker than that of the standard mitochondrial protein-coding genes), is transcriptionally active and undergoes posttranscriptional processing, suggesting that this gene is functional (Boer & Gray, 1988; Popescu & Lee, 2007; Smith & Lee, 2008a; Salinas-Giégé et al., 2017; Gallaher et al., 2018). Its exact role remains unknown.

Vahrenholz et al. (1993) proposed that rtl is involved in C. reinhardtii mtDNA replication, which if true would be unprecedented for an organelle genome. Their hypothesis was based on the fact that an 86-nt stretch of internal mtDNA matches perfectly to the extreme termini of the linear genome, including the 3’ overhangs (Fig. 1a), that is the outermost 86 nt of the two telomeres is repeated once within the mtDNA. This short internal telomeric repeat (or part of it), they suggested, could be transcribed and the resulting RNA used to prime DNA synthesis at the ends of the genome, giving rise to two short internal telomeric repeat (or part of it), they suggested, could be transcribed and the resulting RNA used to prime DNA synthesis on both parental strands (Figs 1b, S1).

The idea that rtl is involved in mtDNA replication is still a compelling one. From a comparative genomics perspective, it would be useful to have mtDNA sequences from very close relatives of C. reinhardtii to see if certain features are conserved such as an internal telomeric repeat. Surprisingly, these kinds of data are lacking. Only a partial mitogenome sequence is available for Chlamydomonas incerta (Popescu & Lee, 2007), the nearest known noninterfertile relative of C. reinhardtii (Schlösser et al., 1976; Pröschold et al., 2001), and no mtDNA sequences are available for either Chlamydomonas schloeseri, the closest identified sister lineage to the C. reinhardtii–C. incerta clade (Pröschold et al., 2018; Nakada et al., 2019) or from strains of Chlamydomonas debaryana (also known as Edaphochlamys debaryana) that consistently group close to C. reinhardtii in phylogenetic analyses (Yumoto et al., 2013; Pröschold et al., 2018; Nakada et al., 2019; Craig et al., 2020). In fact, the nearest relatives of C. reinhardtii to have their presence of the entire left telomere, including the 3’ overhang, is essential to reach a high level of transformation (Duby et al., 2001; Remacle et al., 2006; Larosa, 2012). Unfortunately, the current C. reinhardtii mitochondrial transformation system does not allow for modification of the right telomere or rtl. Second, detailed transcriptional profiling of the C. reinhardtii mtDNA has revealed that an RNA transcript is generated from the internal telomeric repeat, specifically from the antisense strand (Salinas-Giégé et al., 2017), meaning the transcript has the correct polarity to bind to the 3’ overhangs and act as a primer for DNA synthesis on both parental strands (Figs 1b, S1).

Recent years have seen the sequencing of mtDNAs from dozens of other chlamydomonadalean species. Most of these genomes are circular mapping and lack rtl (Smith et al., 2013; Hu et al., 2019; Zhang et al., 2019), but some are linear. The colonial alga Yamagishiella unicocca, for instance, has a linear mitogenome with inverted-repeat telomeres, short internal telomeric repeats and rtl, but the structure of the chromosome ends (e.g. 3’ overhangs) is undetermined (Hamaji et al., 2017). Eudorina sp. NIES-3934 has a unique relative of Y. unicocca and both algae have identical mitochondrial gene orders. Eudorina sp., however, has a primarily circularly mapping genome and, most importantly, is missing rtl (Hamaji et al., 2017), supporting the notion that this gene is involved in mitochondrial telomere maintenance. The only other well-studied example of linear mtDNAs from the Chlamydomonadales is Polytomella. All known members of this nonphotosynthetic genus harbour linear mitogenomes, but they do not contain rtl or any internal copies of the telomeric repeats (Smith et al., 2010). Polytomella mitogenomes terminate in single-stranded hairpins (Smith & Lee, 2008b), meaning that mtDNA replication could easily proceed via a single-stranded circular intermediate and, therefore, is not necessarily dependent on a reverse transcriptase unlike the rtl-based replication model for C. reinhardtii by Vahrenholz and colleagues.

The idea that rtl is involved in mtDNA replication is still a compelling one. From a comparative genomics perspective, it would be useful to have mtDNA sequences from very close relatives of C. reinhardtii to see if certain features are conserved such as an internal telomeric repeat. Surprisingly, these kinds of data are lacking. Only a partial mitogenome sequence is available for Chlamydomonas incerta (Popescu & Lee, 2007), the nearest known noninterfertile relative of C. reinhardtii (Schlösser et al., 1976; Pröschold et al., 2001), and no mtDNA sequences are available for either Chlamydomonas schloeseri, the closest identified sister lineage to the C. reinhardtii–C. incerta clade (Pröschold et al., 2018; Nakada et al., 2019) or from strains of Chlamydomonas debaryana (also known as Edaphochlamys debaryana) that consistently group close to C. reinhardtii in phylogenetic analyses (Yumoto et al., 2013; Pröschold et al., 2018; Nakada et al., 2019; Craig et al., 2020). In fact, the nearest relatives of C. reinhardtii to have their
Mitochondrial genomes from the closest known relatives of *C. reinhardtii*

By mining data from various ongoing nuclear genome projects and publicly available sequencing repositories (Hirashima et al., 2016; Nelson et al., 2019; Craig et al., 2020) (Methods S1), we collected complete or near-complete mtDNA sequences from five different *Chlamydomonas* algae (Fig. 1c). *C. incerta* SAG 7.73, *C. schloesseri* CCAP 11/173 and *C. debaryana* strains NIES-2212, CCAP 11/70 and WS7. The mtDNAs from all of these isolates assemble as linear molecules (17.6–23.1 kb) with long inverted-repeat telomeres (505–1293 nt). Their gene contents and arrangements mirror that of *C. reinhardtii*, including the presence and position of rtl (Fig. 1a). The rtl genes varied in size from 1119–1179 nt, contained no internal stop codons or frameshift mutations and were not associated with any intronic sequences. A multiple alignment of the deduced amino acid sequences of rtl, including that of *C. reinhardtii*, gave an average pairwise-positive identity of 67% (BLSM2 scoring matrix), which went up significantly within the region containing the core reverse transcriptase domain (Fig. S2). In short, rtl appears to be a functional gene in not only *C. reinhardtii* but in its closest known relatives with linear mtDNAs. But is it involved in telomere maintenance?

The most striking finding from these newly sequenced linear mtDNAs is that, like *C. reinhardtii*, the extreme ends of the telomeres are repeated once within the genome (Fig. 1c). Equally as striking is that this internal repeat occurs at approximately the same location and arrangement in all explored strains: within the final 400 nt before the start of the right telomere and with an orientation matching that of the left telomere (and the reverse complement of the right telomere) (Fig. 1c). The length of this internal telomeric repeat ranged from 33 nt to 98 nt. In *C. reinhardtii*, *C. incerta* and *C. schloesseri* it occurred downstream of the rnl2 coding module, whereas in the three *C. debaryana* strains it was within the rnl1 coding module, which is adjacent to rnl2 (Fig. 1c). Consequently, the sequences of these internal repeats can only be aligned among subsets of these six isolates (Fig. 1c). Also, it is noteworthy that the telomeric repeats themselves are highly divergent among the different strains, even between the most closely related ones. The
C. reinhardtii and C. incerta mitochondrial telomeres, for example, share only c. 65% nucleotide identity, not including gaps. This is significant because it means that it is not a specific telomeric sequence that is being maintained in these linear mtDNAs – a common feature of linear organelle genomes (Smith & Keeling, 2013) – but rather a specific genomic architecture, one in which the ends of the genome match to an internal site and one that likely existed in the most recent common ancestor of the core Reinhardtinia clade (Hamaji et al., 2017).

Recurring themes in support of Vahrenholz and colleagues

Although we are confident that near-complete telomeric sequences were assembled from the Chlamydomonas mtDNAs presented here (Methods S1), detailed laboratory work is needed to determine the precise structure of the ends and if they are represented by 3’ overhangs. Nevertheless, a recurring theme arose from these data: linear mtDNAs within C. reinhardtii and its close relatives have a free-standing rtl gene (located between the L3 and L8 rRNA-coding modules) and an internal repeat identical to the extreme termini. This internal repeat is consistently located adjacent to the right telomere with an orientation by which its putative transcriptional product can bind to the 3’ ends of the mtDNA, initiating DNA synthesis. All of this evidence supports the rtl-based mtDNA replication model of Vahrenholz and colleagues.

There are two other competing models for mtDNA replication in C. reinhardtii, both based on the presence of the internal telomeric repeat but not requiring the existence of a reverse transcriptase. One involves the formation of a T-loop through the interaction of the left telomere and internal repeat (Larosa, 2012) (Fig. S3). Indeed, rare circular forms of the C. reinhardtii mtDNA, potentially generated by T-loop formation, have been identified (Ryan et al., 1978; Ma et al., 1992; Duby et al., 2001). However, the location of the internal repeat in the three C. debaryana strains did not support this mechanism, as it would presumably exclude replication of the rnl2 coding module (Fig. S3). Similar to Larosa (2012), Vahrenholz et al. (1993) presented an alternative DNA-based replication model to their rtl-dependent model in which folding of mtDNA provided an interaction between the terminal and internal telomeric repeats, which then primed DNA synthesis. A site-specific endonuclease was required in both of these models, but a free-standing gene encoding such an enzyme was not found in the mtDNA of C. reinhardtii or its close relatives, nor has a nuclear-encoded, mitochondrial-targeted endonuclease been identified. That said, site-specific endonucleases are encoded in the mtDNA group I introns of C. incerta, C. schloesseri and the three C. debaryana strains as well as in certain natural isolates of C. reinhardtii (Smith & Lee, 2008a).

There currently exists an exception to the rule that rtl only occurs in linear mtDNAs. The four-celled colonial chlamydomonadalean Tetraebaena socialis has a circular-mapping mitochondrial genome with a free-standing rtl gene, situated downstream of the rnl3 coding module (the same location as in the linear mtDNAs) (Featherston et al., 2016). The length and deduced amino acid sequence of rtl from T. socialis were similar to that of its counterparts with linear mitogenomes (Fig. S2). We re-assembled de novo the T. socialis mtDNA using the raw sequencing data of Featherston et al. (2016) and our results were consistent with that of a circular map. Given that the transition from a linear to a circular mitogenome architecture is likely to have occurred multiple times independently within the Chlamydomonadales (Hamaji et al., 2017), it is possible that T. socialis recently descended from an ancestor with a linear mitogenome and an rtl-dependent mtDNA replication system. Mapping of T. socialis RNA-sequencing data (GenBank accession no. SRX3567144) to its mtDNA showed very low and limited coverage for this gene, unlike for other mitochondrial protein-coding genes; this could indicate that it is no longer essential and might soon be lost.

Therefore, the mystery of rtl and its function in green algal mtDNAs continues. A clear path to potentially solving this riddle would be mtDNA editing of rtl and the internal telomeric repeat. C. reinhardtii is one of the few species for which an established mitochondrial transformation system exists (Remacle et al., 2006), but it is still not possible to directly target rtl or the internal repeat (Larosa & Remacle, 2013). When such experiments are possible, which might be soon, the findings could be paradigm shifting.

Acknowledgements

DRS is funded by a Discovery Grant from NSERC. RJC is supported by a BBSRC EASTBIO Doctoral Training Partnership.

Author contributions

DRS wrote the initial draft of the manuscript. RJC helped assemble the mitochondrial genomes. All authors contributed to writing and revising the manuscript.

ORCID

Rory J. Craig 1 https://orcid.org/0000-0002-6262-0008
David Roy Smith 2 https://orcid.org/0000-0001-9560-5210

David Roy Smith 1* and Rory J. Craig 2

1Department of Biology, University of Western Ontario, London, ON N6A 5B7, Canada;
2Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh, EH9 3FL, UK
(*Author for correspondence: email dsmit242@uwo.ca)

References

Boer PH, Gray MW. 1988. Genes encoding a subunit of respiratory NADH dehydrogenase (ND1) and a reverse transcriptase-like protein (RTL) are linked to ribosomal RNA gene pieces in Chlamydomonas reinhardtii mtDNA. EMBO Journal 7: 3501–8.


Smith DR, Lee RW. 2008a. Nucleotide diversity in the mitochondrial and nuclear compartments of Chlamydomonas reinhardtii: investigating the origins of genome architecture. BMC Evolutionary Biology. 8: 156.


Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 A hypothetical model for reverse-transcriptase-dependent mtDNA replication in C. reinhardtii and its close relatives, based on Vahrenholz et al. (1993).

Fig. S2 Multiple alignment of the deduced amino acid sequence of rrl from chlamydomonadalean algae.

Fig. S3 Mitochondrial DNA replication via T-loop formation, based on Larosa (2012).

Methods S1 Genomic data used to mine or assemble mtDNAs from Chlamydomonas species.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the New Phytologist Central Office.

Key words: inverted repeat, linear mitochondrial genome, reverse transcriptase, telomere, telomerase.

Received, 17 June 2020; accepted, 4 September 2020.