

# Evolution of linear mitochondrial DNA in three known lineages of *Polytomella*

David Roy Smith · Jimeng Hua · Robert W. Lee

Received: 28 April 2010/Revised: 2 June 2010/Accepted: 4 June 2010/Published online: 24 June 2010  
© Springer-Verlag 2010

**Abstract** Although DNA sequences of linear mitochondrial genomes are available for a wide variety of species, sequence and conformational data from the extreme ends of these molecules (i.e., the telomeres) are limited. Data on the telomeres is important because it can provide insights into how linear genomes overcome the end-replication problem. This study explores the evolution of linear mitochondrial DNAs (mtDNAs) in the green-algal genus *Polytomella* (Chlorophyceae, Chlorophyta), the members of which are non-photosynthetic. Earlier works analyzed the linear and linear-fragmented mitochondrial genomes of *Polytomella capuana* and *Polytomella parva*. Here we present the mtDNA sequence for *Polytomella* strain SAG 63-10 [also known as *Polytomella piriformis* (Pringsheim 1963)], which is the only known representative of a mostly unexplored *Polytomella* lineage. We show that the *P. piriformis* mtDNA is made up of two linear fragments of 13 and 3 kb. The telomeric sequences of the large and small fragments are terminally inverted, and appear to end

in vitro with either closed (hairpin-loop) or open (nicked-loop) structures as also shown here for *P. parva* and shown earlier for *P. capuana*. The structure of the *P. piriformis* mtDNA is more similar to that of *P. parva*, which is also fragmented, than to that of *P. capuana*, which is contained in a single chromosome. Phylogenetic analyses reveal high substitution rates in the mtDNA of all three *Polytomella* species relative to other chlamydomonadalean algae. These elevated rates could be the result of a greater number of vegetative cell divisions and/or small population sizes in *Polytomella* species as compared with other chlamydomonadalean algae.

**Keywords** *Polytomella piriformis* · SAG 63-10 · *Chlamydomonas reinhardtii* · *Volvox carteri* · Telomere

## Introduction

Although it was once assumed that all mitochondrial genomes are circular molecules (or at least circularly mapping molecules), it is now well established that linear and linear-fragmented mitochondrial genomes have evolved numerous times in diverse eukaryotic lineages (for examples and references see Table 1). Linear mitochondrial DNAs (mtDNAs) are defined not only by the fact that they migrate as linear molecules in gel-electrophoresis analyses but also by the presence of specialized terminal structures called telomeres (Nosek and Tomáška 2002, 2003). Nucleotide sequence data of linear mitochondrial genomes are available for a wide variety of species (reviewed by Nosek et al. 2004a); however, in many cases sequence data from the extreme ends (i.e., telomeres) of linear mtDNAs are unavailable. This is primarily because the telomeres of linear mitochondrial genomes often have terminal

Communicated by L. Tomaska.

Sequence data from this article have been deposited in GenBank under accession numbers GU108480 and GU108481.

**Electronic supplementary material** The online version of this article (doi:10.1007/s00294-010-0311-5) contains supplementary material, which is available to authorized users.

D. R. Smith (✉) · J. Hua · R. W. Lee  
Department of Biology, Dalhousie University,  
Halifax, NS, Canada  
e-mail: smithdr@dal.ca

J. Hua  
e-mail: j.hua@dal.ca

R. W. Lee  
e-mail: robert.lee@dal.ca

**Table 1** Available organelle-genome data for chlamydomonadalean green algae (with some examples of non-green-algal linear mtDNAs)

Genus and species	Clade <sup>a</sup> (or group)	Mitochondrial genome			Mitochondrial-DNA telomeres			Reference
		Conformation	Size (kb)	%GC	Conformation of ends	Size (kb)	Telomere type	
<b>Chlamydomonadales<sup>a</sup></b>								
<i>Chlamydomonas asymmetrica</i>	<i>Reinhardtina</i>	Linear	~14	–	–	–	–	Lafamme and Lee (2003)
<i>Chlamydomonas incerta</i>	<i>Reinhardtina</i>	Linear	~17.5 <sup>c</sup>	~44 <sup>c</sup>	–	–	–	Popescu and Lee (2007)
<i>Chlamydomonas reinhardtii</i>	<i>Reinhardtina</i>	Linear	15.8 ⇔ 18.9 <sup>b</sup>	~45	3' extension	~0.6	TIR	Michaelis et al. (1990) Smith and Lee (2008a)
<i>Chlamydomonas sphaeroides</i>	<i>Reinhardtina</i>	Linear	~25f	–	–	–	–	Lafamme and Lee (2003)
<i>Pandorina morum</i>	<i>Reinhardtina</i>	Linear	19 ⇔ 38	–	–	~1.8–3.3	TIR	Moore and Coleman (1989)
<i>Polytomella capuana</i>	<i>Reinhardtina</i>	Linear	13.0	57.2	Hairpin loop or nicked loop	~0.9	TIR	Smith and Lee (2008b)
<i>Polytomella parva</i>	<i>Reinhardtina</i>	Linear bipartite	~13 and 3	41.0 <sup>d</sup>	Hairpin loop or nicked loop	~1.3	TIR	Fan and Lee (2002)
<i>Polytomella piriformis</i>	<i>Reinhardtina</i>	Linear bipartite	~13 and 3	42.0 <sup>d</sup>	Hairpin loop or nicked loop	~1.3	TIR	This study
<i>Sphaerellopsis aulata</i>	<i>Reinhardtina</i>	Linear	~25 <sup>f</sup>	–	–	–	–	Lafamme and Lee (2003)
<i>Volvox carteri f. nagariensis</i>	<i>Reinhardtina</i>	Circular <sup>e</sup>	~35 <sup>c</sup>	~34 <sup>c</sup>	N/A	N/A	N/A	Smith and Lee (2009) Smith and Lee (2010)
<i>Chlamydomonas eugametos</i>	<i>Moewusinia</i>	Circular	22.9	34.6	N/A	N/A	N/A	Denovan-Wright et al. (1998)
<i>Chlamydomonas moewusii</i>	<i>Moewusinia</i>	Circular	~21	–	N/A	N/A	N/A	Boudreau et al. (1994)
<i>Chlamydomonas pilschmannii</i>	<i>Moewusinia</i>	Circular	~16.5	–	N/A	N/A	N/A	Boudreau and Turmel (1995)
<i>Lobochlamys cullenis</i>	<i>Oogamochlamydia</i>	Linear fragmented	–	~60	–	–	–	Borza et al. (2009)
<i>Lobochlamys segnis</i>	<i>Oogamochlamydia</i>	Fragmented <sup>g</sup>	3 ⇔ 20	~55	–	–	–	Borza et al. (2009)
<i>Oogamochlamys gigantea</i>	<i>Oogamochlamydia</i>	Fragmented <sup>g</sup>	–	~50	–	–	–	Borza et al. (2009)
<i>Chlorogonium elongatum</i>	<i>Chlorogonia</i>	Circular	22.7	37.8	N/A	N/A	N/A	Kroymann and Zetsche (1998)
<i>Dunaliella salina</i>	<i>Dunaliellina</i>	Circular	28.3	34.4	N/A	N/A	N/A	Smith et al. (2010)
<b>Examples of other linear mtDNAs</b>								
<i>Aurelia aurita</i>	Cnidaria	Linear	16.9	33.3	–	~0.4	TIR	Shao et al. (2006)
<i>Hydra magnipapillata</i>	Cnidaria	Linear bipartite	8.2 and 7.7	23.7 <sup>d</sup>	–	~0.2	TIR	Voigt et al. (2008)
<i>Hydra oligactis</i>	Cnidaria	Linear	16.3	23.7	–	~1.5	TIR	Kayal and Lavrov (2008)
<i>Candida parapsilosis</i>	Yeast	Linear	32.7	23.8	5' extension	~1.4	TIR (TR)	Nosek et al. (2004b)
<i>Ochromonas danica</i>	Chrysoophyta	Linear	41.0	~26	N/A	N/A	N/A	Coleman et al. (1991)
<i>Tetrahymena pyriformis</i>	Ciliata	Linear	42.1	~21	–	2.7	TIR (TR)	Burger et al. (2000)

**Table 1** continued

Genus and species	Clade <sup>a</sup> (or group)	Mitochondrial genome		Mitochondrial-DNA telomers			Reference		
		Conformation	Size (kb)	%GC	Conformation of ends	Size (kb)		Telomere type	%GC
<i>Theileria parva</i>	Apicomplexa	Linear	6.6	21.3	–	~0.45	TIR	~29	Shukla and Nene (1998)

N/A not applicable, – data not available, *TIR* terminal inverted repeat, *TIR* (TR) terminal inverted repeat containing internal tandem repeats

<sup>a</sup> Our definition of Chlamydomonadales follows that of Lewis and McCourt (2004); clades are defined by Nakada et al. (2008)

<sup>b</sup> These data vary because of the presence/absence of optional introns (Smith and Lee 2008a)

<sup>c</sup> These data are based on almost-complete genome sequences

<sup>d</sup> MtDNA consists of two fragments; data are based on the concatenation of these fragments

<sup>e</sup> The circular conformation of the *V. carteri* mtDNA is based on genome-assembly data (Smith and Lee 2010) and needs to be confirmed by gel-electrophoresis results

<sup>f</sup> MtDNAs were at the limit of the linear size range for the gels

<sup>g</sup> It is unknown, at present, if these fragments have circular or linear conformations (Borza et al. 2009)

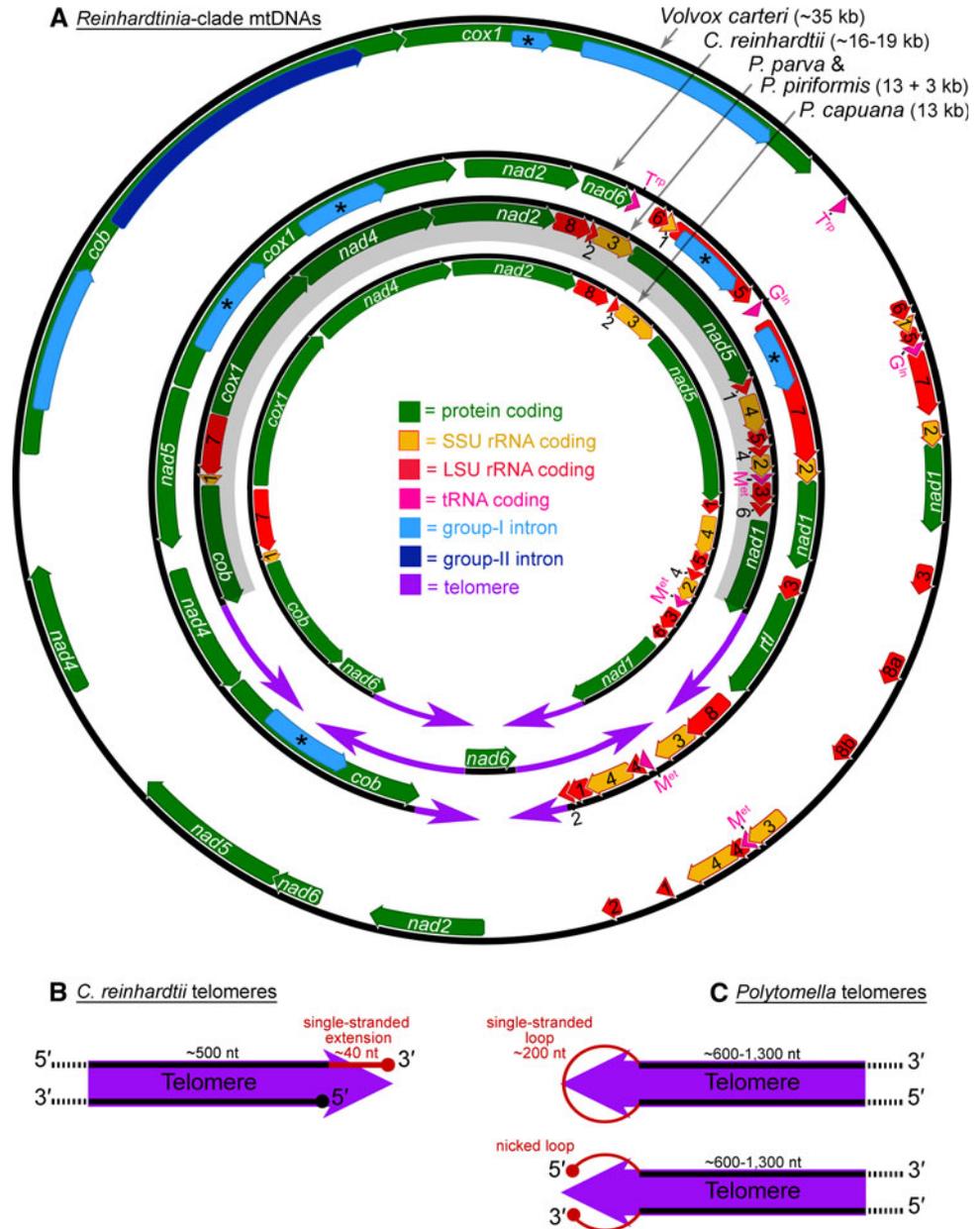
conformations that are unamenable to PCR and cloning, such as 3' or 5' overhangs or covalently closed ends (Nosek and Tomáška 2002, 2003). Indeed, many of the linear mtDNAs that are listed in GenBank as “complete genome sequences” are missing sequence data from their telomeres. This is unfortunate because by knowing the sequence and end conformation of telomeres, one can gain insights into how linear mitochondrial genomes overcome the end-replication problem, as defined by Olovnikov (1971) and Watson (1972), a problem which is not only interesting from a genome-evolution perspective but one which has important implications for cancer research (Nosek et al. 2004a).

One eukaryotic lineage that has many species with linear mtDNAs is the “*Reinhardtinia* clade,” *sensu* Nakada et al. (2008), a monophyletic group of green algae found within the chlorophycean class of the Chlorophyta (Lewis and McCourt 2004). All of the *Reinhardtinia*-clade mtDNAs that have been examined appear to be either linear or linear-fragmented (Laflamme and Lee 2003), with the exception of the *Volvox carteri* mtDNA, which assembles as a circular molecule (Smith and Lee 2010). Conversely, all of the characterized green-algal mitochondrial genomes from outside this clade map as circular molecules, save for the mtDNAs of some *Lobochlamys* taxa, which may be linear-fragmented (Borza et al. 2009).

The first mtDNA from the *Reinhardtinia* clade to be completely sequenced was that of the model organism *Chlamydomonas reinhardtii* (Gray and Boer 1988; Michaelis et al. 1990; Smith and Lee 2008a). The telomeres of this 16–19 kb linear genome are each composed of an ~600 nucleotide (nt) inverted repeat with an ~40 nt, non-complementary 3' extension (Fig. 1a, b); based on the potential interaction of the telomeres with internal repeats, two models of mitochondrial-genome replication have been proposed (Vahrenholz et al. 1993). The only other well-studied *Reinhardtii*-clade linear mtDNAs come from algae within the *Polytomella* genus—a group of non-photosynthetic, naturally wall-less unicells (Pringsheim 1955).

Fan and Lee (2002) sequenced the *Polytomella parva* mtDNA and reported that it is made up of two linear fragments of ~13.5 and ~3.5 kb with nine and one gene, respectively (Fig. 1a). Like the *C. reinhardtii* mtDNA, the telomeres of both fragments form ~1.3 kb inverted repeats; however, it was assumed that the DNA sequence of the extreme telomere ends was not determined and no potential telomere conformation was proposed. Further work by Mallet and Lee (2006), using a *cox1* phylogeny, revealed that there are at least three distinct lineages of *Polytomella*: two that are closely related and represented by *P. parva* and *Polytomella* strain SAG 63-10 [referred to hereafter by its original *nomen nudum* name of *Polytomella piriformis* (Pringsheim 1963)], and a third, deep-branching

**Fig. 1** Genome maps of completely sequenced *Reinhardtinia*-clade mtDNAs. **a** Arrows within the coding regions indicate transcriptional polarities. The SSU- and LSU-rRNA-coding modules (i.e., *rrns*- and *rrnl*-gene fragments) are numbered based on their order in the gene. Portions of the *P. parva* and *P. piriformis* mitochondrial-genome maps that are shaded gray represent regions of gene colinearity with the *P. capuana* mtDNA. Introns annotated with an asterisk are optional—see Smith and Lee (2008a, 2009) for a list of the *C. reinhardtii* and *V. carteri* strains in which the different introns occur. **b**, **c** In vitro mtDNA telomere conformation for *C. reinhardtii* and *Polytomella* taxa. Sequencing and agarose-gel electrophoresis analyses suggest that the *Polytomella* telomeres can exist in either a closed (hairpin-loop) or an open (nicked-loop) conformation (see Online Resource 1, Supplementary Figure S1A for more details)



lineage represented by *Polytomella capuana*. Gel electrophoresis and Southern blot analyses suggested that the *P. capuana* mtDNA is contained in a single linear chromosome (Mallet and Lee 2006) whereas unpublished data (Mallet MA and Lee RW, personal communication) seemed to indicate that the mitochondrial genome of *P. piriformis*, like that of *P. parva*, is a linear bipartite molecule.

Smith and Lee (2008b) confirmed the linear monomeric conformation of the *P. capuana* mitochondrial genome by obtaining its complete DNA sequence. They found the *P. capuana* mtDNA to be a 13-kb linear molecule with ~0.9 kb inverted-repeat telomeres that terminate in vitro with either a closed (hairpin-loop) or an open (nicked-loop) conformation (Fig. 1a, c). It was also shown that the gene

located on the ~3.5 kb mtDNA chromosome in *P. parva* (i.e., *nad6*) is found internal to the left telomere on the *P. capuana* mtDNA; based on the position of this gene, Smith and Lee (2008b) proposed a recombination-driven scenario of how the linear intact mtDNA of *P. capuana* could be converted into a linear-fragmented form like that of *P. parva*. An additional and unexpected observation was that the *P. capuana* mtDNA, unlike that of *P. parva*, is rich in guanine and cytosine (GC rich), which is an unusual and, at that time, unprecedented trait for an organelle genome (Smith and Lee 2008b).

This study examines mitochondrial genome evolution in *P. piriformis*, the only known representative of a third and mostly unexplored *Polytomella* lineage (Mallet and Lee 2006).

The mtDNA sequence of *P. piriformis*, including the sequence and structure of the telomeres, is described and compared with those from other *Polytomella* species. Moreover, to complement our *P. piriformis* mtDNA sequence data, we re-examined the *P. parva* mtDNA telomeres, which were reported to be only partially sequenced (Fan and Lee 2002). The ultimate aims of this study are to: (1) confirm if the mtDNA of *P. piriformis* is fragmented like that of *P. parva* or linear-intact like that of *P. capuana*; (2) gain insight into mtDNA-telomere evolution and how linear mitochondrial genomes overcome the end-replication problem; and (3) provide complete mtDNA sequence data, including the structure of the telomeres, for members from each of the three known *Polytomella* lineages.

## Methods

### Culture conditions, mitochondrial enrichment, and DNA extraction

*Polytomella* strain SAG 63-10 was obtained in 2004 from the Sammlung von Algenkulturen in Göttingen, Germany, whereas *P. parva* (UTEX L 193) was obtained from the Culture Collection of Algae at the University of Texas at Austin around 1994. SAG 63-10 was originally isolated by E.G. Pringsheim at the Göttingen Botanical Gardens sometime before 1963 [Pringsheim 1963; SAG website (<http://sagdb.uni-goettingen.de/>)]. When Pringsheim first described strain 63-10 he named it *Polytomella piriformis* (Pringsheim 1963), a name we will use here; however, this name is currently considered a *nomen nudum* because a taxonomically valid description of *P. piriformis* was never published.

*Polytomella piriformis* and *P. parva* were grown at 22°C in *Polytomella* medium (0.1% tryptone, 0.2% yeast extract, 0.2% sodium acetate trihydrate) and harvested in the late logarithmic phase of growth ( $OD_{750\text{ nm}} = 0.4$ ; determined with a Bausch and Lomb Spectronic 20 spectrophotometer) by centrifugation (1,000g) at 4°C. Disruption of cells was carried out using a Dounce homogenizer. A mitochondria-enriched fraction was obtained by following Procedure B of Ryan et al. (1978). Isolation of DNA followed the method of Ryan et al. (1978), with the exception that there was no further DNA purification step employing preparative CsCl gradient centrifugation.

### Amplification, digestion, and sequencing of DNA

PCR experiments were performed in High Fidelity Platinum SuperMix (Invitrogen, Carlsbad, CA) using DNA from a mitochondria-enriched fraction as the template. The telomeric regions of the *P. piriformis* mtDNA were

amplified using long-walk PCR (Katz et al. 2000) and terminal deoxynucleotidyl transferase (TdT) tailing (Förstemann et al. 2000; Bah et al. 2004). PCR-, long-walk PCR-, and TdT-tailing-products were separated by agarose-gel electrophoresis, purified with the QIAquick Gel Extraction Kit (Qiagen, Germantown, MD), then cloned using the TOPO TA Cloning Kit (Invitrogen). Plasmid DNA was extracted with the QIAquick Spin Miniprep Kit (Qiagen). Reverse-transcriptase (RT) PCR reactions were performed with the SuperScript III One-Step RT-PCR System (Invitrogen) following the manufacturer's protocol. Purified PCR products and isolated plasmids were sequenced on both strands using a 3730xl DNA Analyzer (Applied Biosystems, Foster City, CA) at the MacroGen Sequencing Facility in Rockville, MD. For restriction enzyme reactions, DNA from a mitochondria-enriched fraction was treated with *Xba*I and *Xho*I (10 U of each enzyme, ~2 µg mtDNA, RE<sub>ACT</sub> buffer 2; Invitrogen) for 2 h at 37°C.

### DNA blotting and hybridization

Blotting of agarose gels onto Hybond N+ membranes (GE Healthcare, Buckinghamshire, UK) was performed using the Vacublot XL System (GE Healthcare). Probes used in this study were labeled and hybridized to samples with the AlkPhos Direct Labeling and Detection System (GE Healthcare) following the manufacturer's instructions. Label was detected by exposing the membranes to Super RX medical X-ray film (Fuji Photo Film Co., Tokyo, Japan).

### Sequence analysis

Sequences were assembled using CodonCode Aligner Version 3.0.1 (CodonCode Corporation, Dedham, MA), which employs the Phred, Cross-match, and Phrap algorithms for base calling, sequence comparison, and sequence assembly, respectively. Assemblies were performed with a minimum-percent-identity score of 98, a minimum-overlap length of 500 nt, a match score of 1, a mismatch penalty of -2, a gap penalty of -2, and an additional first-gap penalty of -3. The boundaries of the mitochondrial rRNA-coding modules were estimated by sequence comparisons with their counterparts in the *P. parva* and *P. capuana* mtDNAs. Multiple alignments of DNA sequences were performed with CLUSTAL W (Thompson et al. 1994) using the default settings. The maximum-likelihood phylogeny was constructed with PAUP\* version 4.0b10 (Swofford 2003) using the general time reversible model (GTR + I + G), which was chosen by Modeltest [version 3.7 (Posada and Crandall 1998)] as the best-fit model of nucleotide substitution for our given

dataset. The tree was built using the first- and second-codon positions of the seven mtDNA-encoded proteins that are shared among all completely sequenced chlorophycean mtDNAs.

#### Nucleotide sequence accession numbers

The nucleotide sequence of the *P. piriformis* mitochondrial genome is deposited in GenBank under the accession numbers GU108480 (13-kb fragment) and GU108481 (3-kb fragment).

## Results

### *Polytomella parva* mtDNA telomeres

To complement our *P. piriformis* mtDNA sequence data, we attempted to finish sequencing the *P. parva* mtDNA telomeres. Fan and Lee (2002), using agarose-gel-electrophoresis data, estimated that the *P. parva* mitochondrial genome is separated into two fragments of 13.5 and 3.5 kb. Thus, when Fan and Lee (2002) cloned, sequenced, and assembled the two *P. parva* mtDNA fragments and ended up with contigs of 13 and 3 kb, they believed that they were missing approximately 250 nt of telomeric sequence from the ends of each chromosome. Here, using DNA isolated from a mitochondria-enriched fraction of *P. parva* (UTEX L 193), we performed TdT-tailing experiments (Förstemann et al. 2000; Bah et al. 2004) on the large and small mtDNA fragments. Analyses of the resulting telomeric sequence data indicate that Fan and Lee (2002) completely sequenced the *P. parva* mitochondrial genome—we found no evidence that the telomeres extend beyond the sequence that they generated. Therefore, the two *P. parva* mtDNA fragments appear to have lengths of ~13 and 3 kb. The reason why these mtDNAs migrate as ~13.5- and 3.5-kb molecules in agarose-gel-electrophoresis analyses may be linked to the conformation of their telomeres (discussed further below).

### Characterization of the *P. piriformis* mitochondrial genome

When DNA was isolated from a mitochondria-enriched fraction of *P. piriformis* and run on agarose gels, two distinct bands with estimated sizes of 13.5 and 3.5 kb (referred to hereafter as the large and small fragments) were consistently observed after staining with ethidium bromide (Online Resource 1, Supplementary Figure S1). Both of these bands co-migrated with linear-DNA markers under various concentrations of agarose (data not shown), supporting the hypothesis that they represent linear

molecules. DNA sequencing confirmed that both fragments are components of the *P. piriformis* mitochondrial genome. This conclusion was supported by Southern blot hybridization experiments using *P. piriformis*-mtDNA probes (Online Resource 2, Supplementary Figure S2). Southern blot analyses also identified a *P. piriformis*-mtDNA-hybridizing band of ~1.8 kb, which was not visible on ethidium-bromide-stained gels but was consistently observed in Southern blots using different DNA isolates of *P. piriformis* (Online Resource 2, Supplementary Figure S2). Probes corresponding to the *P. piriformis* mtDNA telomeres as well as those coming from the gene found on the small fragment (*nad6*) hybridized to the 1.8-kb component, whereas those designed from genes on the large fragment did not hybridize to this component. For these reasons, we believe that the 1.8-kb component is derived from the small fragment but does not represent a primary constituent of the *P. piriformis* mitochondrial genome. One hypothesis could be that the 1.8-kb piece is a replicative intermediate of the 3-kb fragment; but if this is the case, why are we not able to observe a replicative intermediate of the 13-kb fragment? Another hypothesis could be that the 1.8 kb piece is a distinct mtDNA chromosome that is maintained in addition to the 13- and 3-kb fragments; however, this does not explain why it is found in much smaller concentrations in gel-electrophoresis analyses relative to the two other fragments. Similar findings were observed and discussed for *Polytomella* strains coming from the *P. parva* lineage, where in addition to the two main mtDNA bands, Southern blot analyses identified mtDNA-hybridizing components with sizes of 1.8 and 2.1 kb (Fan and Lee 2002; Mallet and Lee 2006); moreover, *Polytomella papillata* and *Polytomella magna* (both of the *P. parva* lineage) also showed a mtDNA-hybridizing band of 5 kb.

### General features of the *P. piriformis* mitochondrial genome

Complete genetic maps of the *P. piriformis* large- and small-mtDNA fragments are shown in Fig. 1a. For comparison, this figure also contains the mtDNA genetic maps of *P. capuana* (Smith and Lee 2008b), *P. parva* (Fan and Lee 2002), *C. reinhardtii* (Gray and Boer 1988), and *Volvox carteri* (Smith and Lee 2010). Note that unlike the mtDNAs of *P. piriformis* and *P. parva*, which are linear bipartite genomes, those of *P. capuana* and *C. reinhardtii* are linear monomeric molecules, and that of *V. carteri* assembles as a circular molecule (Fig. 1a). The sequenced lengths of the two *P. piriformis* mtDNA fragments are 13,004 and 3,079 nt (accumulative length = 16,083 nt), which is about the same as that of the two *P. parva* mtDNAs (3,018 and 13,135 nt; accumulative length = 16,153 nt)

and around 3-kb larger than the linear monomeric mtDNA of *P. capuana*, which, at 12,998 nt, is the smallest archaeplastidial mitochondrial genome observed to date. As was observed for *P. parva* and *P. capuana*, the sequenced lengths of the *P. piriformis* mtDNAs are approximately 500 nt smaller than what agarose-gel-electrophoresis results would suggest. We believe that conformation of the mtDNA telomeres is retarding migration of these molecules in agarose gels (discussed below).

In addition to DNA sequencing, the sizes of the *P. piriformis* mtDNAs were confirmed by restriction-endonuclease-digestion experiments (Online Resource 1, Supplementary Figure S1). We were unable to PCR amplify circular forms of either the large or small fragments, nor could we recover products that linked the two molecules together.

Altogether, the *P. piriformis* mitochondrial genome [GenBank accession numbers GU108480 (13-kb fragment) and GU108481 (3-kb fragment)] contains ten genes, nine of which are found on the large fragment (representing six proteins, one tRNA, and two rRNAs) and one on the small fragment (*nad6*). The genes on the large fragment are organized into two unequally sized clusters with opposing transcriptional polarities that proceed outwards toward the ends of the chromosome. The sole tRNA-coding gene, *trnM*, has characteristics that are consistent with a role in elongation rather than initiation, as is true for the *trnM* found in other sequenced *Reinhardtinia*-clade mtDNAs. The two ribosomal-RNA-coding genes (*rrn1* and *rrn5*), like those of other chlorophycean mitochondrial genomes, are fragmented and scrambled (Fig. 1a). The degree to which the *rrn1* and *rrn5* regions are fragmented (four and eight coding modules, respectively) is the same as that of the *C. reinhardtii* and *V. carteri* mitochondrial genomes. The mtDNA gene complement and gene arrangement for *P. piriformis* are identical to those of *P. parva* (which also has a bipartite mtDNA with *nad6* isolated on the smallest fragment) and *P. capuana*, except that the *P. capuana* mtDNA is contained in a single chromosome with *nad6* located internal to the left terminal repeat (Fig. 1a). With only ten mtDNA-encoded genes, none of which harbor introns, these three *Polytomella* strains have the most reduced mitochondrial genomes observed from the Archaeplastida. The only species known to have even smaller mtDNA gene repertoires are found in the phyla Apicomplexa and Dinoflagellata, and arguably some species within the supergroup Excavata (Kairo et al. 1994; Nash et al. 2007).

Of the 16,083 nt that comprise the *P. piriformis* mitochondrial genome, 10,590 nt (66%) are coding and 5,493 nt (34%) are noncoding. The latter category can be subdivided into telomeres and intergenic spacers, which represent 5,233 nt (32.5%) and 260 nt (1.6%) of the genome,

respectively. If the telomeres are ignored, the *P. piriformis* mtDNA is only 1.6% noncoding, which is inordinately compact for a green-algal mitochondrial genome (see Smith and Lee 2009 for a compilation). Relative to that of *P. piriformis*, the *P. parva* mtDNA has a comparable distribution of coding, intergenic and telomeric DNA, whereas the *P. capuana* mtDNA has more intergenic DNA (555 vs. 260 nt) and the telomeres constitute a smaller fraction of the genome (13.7 vs. 32.5%); this is because they are shorter than those of *P. piriformis* and *P. parva* (~0.9 vs. ~1.3 kb) and because there are only two of them rather than four as in the *P. piriformis* and *P. parva* mitochondrial genomes. Unlike the *P. capuana* mtDNA, where inverted repeats, which can fold into stem-loop structures, punctuate all but two of the genes, the *P. piriformis* mitochondrial genome contains no obvious inverted-repeat elements in its intergenic regions. Re-analyses of the *P. parva* mtDNA reveal the same conclusion. We cannot say if the lack of short inverted repeats in the *P. piriformis* and *P. parva* mtDNAs is due to an alternative mode of mitochondrial-transcript processing relative to *P. capuana*. However, RT-PCR analyses with either *P. piriformis*, *P. parva*, or *P. capuana* RNA can generate products that span more than one mitochondrial-encoded gene (up to at least five genes), suggesting that polycistronic mitochondrial transcripts are being generated in all three of these algae. We were also able to generate RT-PCR products that spanned most of the telomeric regions from these three *Polytomella* species.

The GC content of the *P. piriformis* mtDNA (42% overall; 41 and 43% for the large and small fragments, respectively) is unremarkable in comparison to that of *P. capuana*, which is one of the more GC-rich mtDNAs observed to date (57%), but comparable to that of *P. parva* (41% overall, 40% large fragment, and 42% small fragment). Among the different regions of the *P. piriformis* mitochondrial genome, the telomeres have a higher GC content (45%) than the coding segments (40%) and the intergenic spacers (37%). This trend is also observed for the *P. parva* mtDNA. The GC contents of the different mtDNA codon positions for *P. piriformis* are 45% (1st position), 39% (2nd position), 36% (3rd position), and 25% (3rd-position synonymous sites), which is similar to that of *P. parva* [44%<sub>(GC1)</sub>, 39%<sub>(GC2)</sub>, 35%<sub>(GC3)</sub>, and 24%<sub>(GCsyn)</sub>] but different than that of *P. capuana* [52%<sub>(GC1)</sub>, 41%<sub>(GC2)</sub>, 76%<sub>(GC3)</sub>, and 85%<sub>(GCsyn)</sub>]. The fact that the nucleotide composition of the *P. piriformis* mtDNA is most biased at what are typically regarded as among the most neutrally evolving positions in a genome (intergenic and synonymous sites) and is least biased at the more functionally constrained first- and second-codon positions, suggests that a neutral process, such as mutation pressure or biased gene conversion, is driving the nucleotide composition towards

A and T. Similarly, it has been argued that a neutral process is biasing the nucleotide content of the *P. capuana* mtDNA; however, unlike for the *P. piriformis* and *P. parva* mtDNAs, this process appears to be skewed towards G and C (Smith and Lee 2008b).

Another notable feature of the *P. piriformis* mtDNA is that the *nad5* gene has an alternative start codon: it uses “GTG” instead of “ATG.” The *nad5* start codon of *P. parva* is also “GTG” but that of *P. capuana* is the canonical “ATG.” Our RT-PCR analyses did not reveal any indication that the *P. piriformis nad5* “GTG” start codon is converted to “AUG” via RNA editing. And although it remains to be confirmed, it is reasonable to assume that the *P. parva nad5* “GTG” start codon also remains unedited in RNA transcripts. These data support the generally held belief that there is no RNA editing in green-algal mitochondria (Nedelcu et al. 2000; Steinhäuser et al. 1999; Turmel et al. 2010). To the best of our knowledge, this is the first recorded example of a chlorophycean mitochondrial genome with an alternative start codon. There are, however, many examples of Metazoan mtDNAs that use “GTG” start codons (<http://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html/index.cgi?chapter=cgencodes>).

#### *P. piriformis* mtDNA telomeres

We obtained the *P. piriformis* mtDNA telomeric sequences using terminal TdT-tailing (Förstemann et al. 2000; Bah et al. 2004) and long-walk PCR (Katz et al. 2000). The former allowed us to amplify the 3′ ends of the telomeres whereas the latter approach amplified the 5′-telomeric ends. See Smith and Lee (2008b) for a more detailed description of how these techniques can be used to amplify telomeric regions.

The *P. piriformis* mitochondrial telomeric sequences form an inverted repeat (i.e., the sequence of one terminus is the reverse complement of the other terminus). Inverted-repeat telomeres are a common feature among species with linear mtDNAs, and are found in *P. parva*, *P. capuana*, *C. reinhardtii*, and *Pandorina morum* as well as in certain apicomplexans, ciliates, cnidarians, and fungi (Table 1). Moreover, linear mitochondrial plasmids typically end in inverted repeats (Handa 2007). Each of the *P. piriformis* mtDNA telomeres is ~1.3 kb in length, which is about the same size as the *P. parva* telomeres and 0.4 kb longer than those of *P. capuana* (Table 1). Within each of the *P. piriformis* mtDNA fragments, the two telomeric sequences are identical, but the telomeres between fragments differ slightly (pairwise identity = 93.7%)—for *P. parva*, alignments of the mtDNA telomeres of the different fragments gives pairwise identity values of ~99.5%. We do not know why there is more sequence divergence between the

telomeric sequences of the two *P. piriformis* mtDNA fragments as compared to those of *P. parva*. One explanation is that the mtDNA of *P. piriformis* has a higher mutation rate and/or a lower rate of intermolecular gene conversion relative to that of *P. parva*. Alignments of the *P. piriformis* telomeres with those of *P. parva* gave pairwise identity values of ~54% (Online Resource 3, Supplementary Figure S3). These alignments reveal long stretches of unrelated sequences punctuated with short blocks of what appears to be conserved sequence elements (Online Resource 3, Supplementary Figure S3). We were unable to align either the *P. piriformis* or *P. parva* mtDNA telomeres with those of *P. capuana* or *C. reinhardtii*.

TdT-tailing and long-walk-PCR analyses suggest that the *P. piriformis* telomeres exist in vitro in two separate conformations: a closed (hairpin-loop) conformation and an open (nicked-loop) conformation (Fig. 1c; Online Resource 1, Supplementary Figure S1). The nucleotide sequences of these two telomeric conformations appear to be identical to one another; however, that from the closed conformation seems to terminate with an ~150-nt single-stranded loop, whereas in the open conformation this loop is nicked. The location of the nick was shown to vary but was most often observed at the apex of the loop. Standard-PCR techniques confirmed the sequence of the telomeric regions. Agarose-gel-electrophoresis and restriction-digest results provide further support for the idea that the *P. piriformis* mtDNA telomeres can have an open or closed conformation (Online Resource 1, Supplementary Figure S1). These findings are consistent with data on the *P. capuana* mtDNA telomeres, which were shown to be either hairpin- or nicked-loop structures, but for *P. capuana* the loop portion of the hairpin is ~220 nt in length. Preliminary analyses (Smith DR and Lee RW, unpublished data) suggest that the mitochondrial telomeres of *P. parva* are also found in either open or closed conformations. It should be emphasized that at present we cannot rule out the possibility that the nicked-loop telomeric conformation observed for the *P. piriformis* mtDNA is the result of nicking during the DNA extraction process. The telomeric conformations may help explain why the *P. piriformis* mtDNA fragments migrate as 13.5- and 3.5-kb molecules in agarose-gel-electrophoresis analyses instead of their sequenced sizes of 13 and 3 kb. One would expect a linear genome with a 150-nt single-stranded hairpin loop at each end (or with ~75 nt of non-complementary DNA at each end—i.e., nicked loops) to migrate slower in an agarose gel than a linear-DNA molecule of the same size but with blunt ends. This same argument applies to the *P. parva* and *P. capuana* mtDNAs, which also migrate in agarose gels at sizes larger than what DNA sequencing would suggest (Mallet and Lee 2006; Smith and Lee 2008b).

## Discussion

### Insights in mitochondrial telomeres

Hairpin-loop telomeres are not uncommon. In addition to *Polytomella* mtDNA, they are found on the mitochondrial genomes of *Pichia*, *Williopsis*, and *Paramecium* (Pritchard and Cummings 1981; Dinouël et al. 1993), the mitochondrial plasmid of the plant pathogenic fungus *Rhizoctonia solani* (Miyashita et al. 1990), the genomes of certain viruses (Baroudy et al. 1983; González et al. 1986; Rohozinski et al. 1989), and the bacterial plasmids of *Borrelia burgdorferi* (Hinnebusch and Barbour 1991). Their wide distribution suggests that hairpin-loop telomeres may provide an effective strategy for overcoming the end-replication problem.

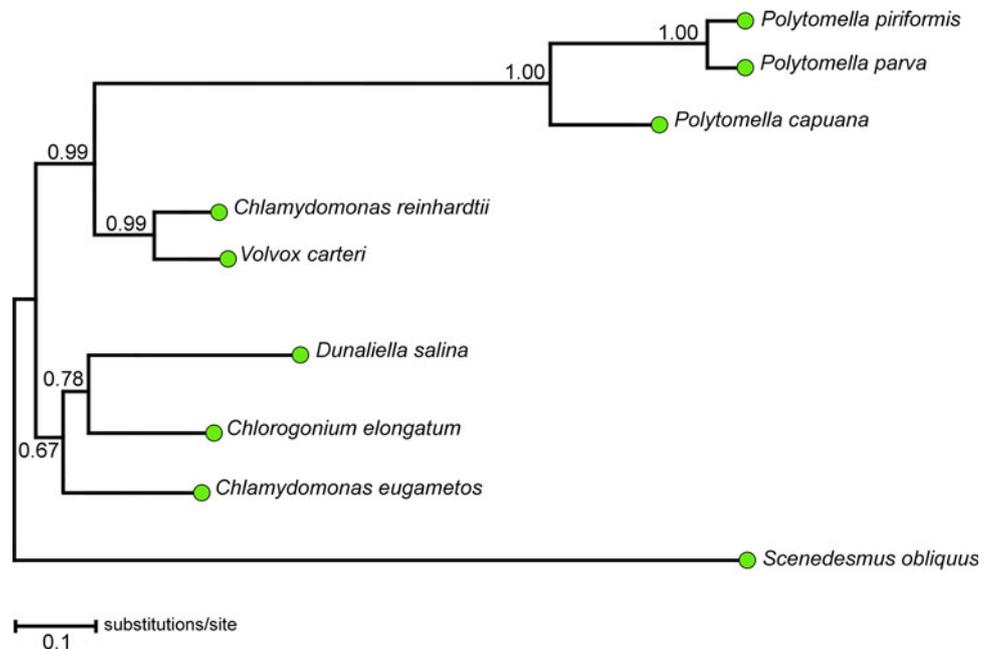
All linear genomes must overcome the end-replication problem, as defined by Olovnikov (1971) and Watson (1972). Unlike eukaryotic nuclear genomes, which mostly employ telomerase to overcome this problem, it is generally accepted that linear mtDNAs do not use telomerase. A variety of telomerase-independent replication models for linear genomes with terminal hairpins have been suggested (Cavalier-Smith 1974; Bateman 1975; Pritchard and Cummings 1981; Baroudy et al. 1983; Dinouël et al. 1993; Traktman 1996). Many of these models center on the fact that a DNA molecule with closed ends is essentially a single-stranded circle, meaning that conventional DNA replication should be able to amplify the entire molecule. Based on the nucleotide sequences of *P. piriformis*, no obvious mode of replication is apparent. However, the fact that we observe both closed and nicked-loop terminal

conformations is indicative of a Cavalier-Smith, Bateman-type model (1975), where conventional DNA replication proceeds through the terminal hairpins and generates a double-stranded circular molecule, which is then nicked at the palindromic sequences by endonucleases followed by unpairing of the two strands and finally self-pairing and ligation of the ends.

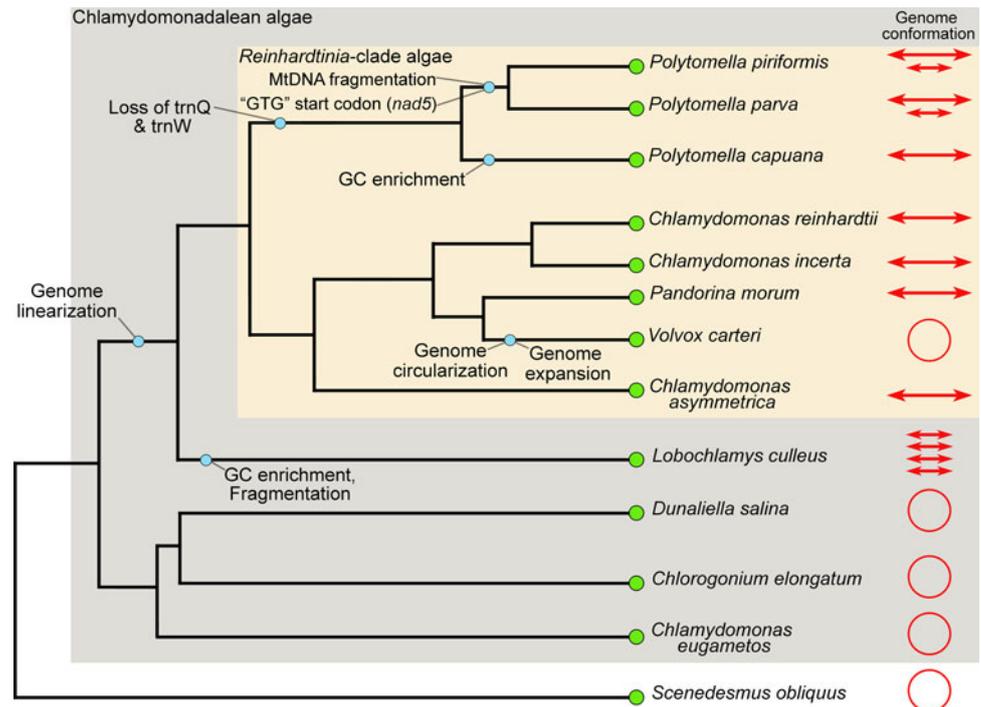
### A *Polytomella* phylogeny

Completion of the *P. piriformis* mitochondrial genome sequence allows for a more thorough phylogenetic evaluation of *Polytomella* species. Previously, Mallet and Lee (2006) performed a phylogenetic analysis using *cox1* sequence data from seven *Polytomella* strains, including *P. piriformis*, and showed that these strains form three distinct lineages, which are represented by *P. capuana*, *P. parva*, and *P. piriformis*, with the *P. parva* lineage containing five of the seven strains tested. Here, we ran a maximum-likelihood phylogeny using the concatenated DNA sequences of seven mtDNA-encoded proteins from the three *Polytomella* species plus six other chlorophycean algae for which complete or almost-complete mitochondrial genome sequences are available. Our results support those of Mallet and Lee (2006): we find that *P. piriformis*, *P. parva*, and *P. capuana* represent separate distinct lineages within the “*Polytomella* clade,” and that *P. capuana* is the deepest branching of the three *Polytomella* species (Fig. 2). This analysis, as well as that of Mallet and Lee (2006), reveals high rates of mtDNA evolution for all three *Polytomella* species relative to other chlamydomonadalean algae (Fig. 2). Previous

**Fig. 2** Maximum-likelihood tree inferred from the DNA sequences of the seven mtDNA-encoded proteins that are shared among all completely sequenced chlorophycean mtDNAs. GenBank accession numbers are as follows: *P. piriformis* (GU108480, GU108481), *P. capuana* (EF645804), *P. parva* (AY062933, AY062934), *C. reinhardtii* (EU306622), *V. carteri* (GU084821), *D. salina* (GQ250045), *C. eugametos* (AF008237), *C. elongatum* (Y13643, Y13644, Y07814), and *S. obliquus* (X17375). All taxa are members of the Chlorophyceae



**Fig. 3** Hypotheses on the evolution of chlamydomonadalean mitochondrial genome architecture. Phylogenetic relationships are based on the 18S-rDNA gene tree of Nakada et al. (2008). Blue circles represent hypotheses on chlamydomonadalean mitochondrial genome evolution. Refer to Table 1 for a complete list of references and GenBank accession numbers



phylogenetic analyses (Pröschold et al. 2001; Nakada et al. 2008) show high rates of nucleotide substitution in the *P. parva* nuclear-encoded SSU rRNA gene relative to that of other chlamydomonadalean algae, therefore, suggesting that a rapid evolutionary rate is a feature of the *Polytomella* lineage and not any one genetic compartment. Such a lineage feature could result from an enhanced number of mutations and/or an elevated level of mutation fixation in the *Polytomella* lineage relative to that of other chlamydomonadalean algae as discussed earlier for *P. parva* (Fan and Lee 2002).

#### Mitochondrial genome evolution of *Reinhardtinia*-clade algae

*P. piriformis* is the seventh chlamydomonadalean alga (fourth from the *Reinhardtinia* clade) to have its mtDNA completely sequenced; moreover, almost-complete mtDNA sequences are available for *V. carteri* and *C. incerta*, and mtDNA gel-electrophoresis results exist for a variety of other chlamydomonadalean species (Table 1). Altogether, these data allow for an intricate picture of the evolution of mitochondrial genome architecture in the Chlamydomonadales. We have summarized on a phylogenetic tree some of the diverse features of mitochondrial genome architecture in this group and proposed the order of events leading to their origin (Fig. 3). It remains to be determined if this diversity in mitochondrial genome architecture in the Chlamydomonadales is greater than that

of other comparable green-algal groups or if it is the result of unequal sampling.

**Acknowledgments** We thank Tudor Borza for insightful discussions during the data analysis. This work was supported by a grant to RWL from the Natural Sciences and Engineering Research Council (NSERC) of Canada. DRS is an Izaak Walton Killam Memorial Scholar and holds a Canada Graduate Scholarship from NSERC.

#### References

- Bah A, Bachand F, Clair E, Autexier C, Wellinger RJ (2004) Humanized telomeres and an attempt to express a functional human telomerase in yeast. *Nucleic Acids Res* 32:1917–1927
- Baroudy BM, Venkatesan S, Moss B (1983) Structure and replication of vaccinia virus telomeres. *Cold Spring Harb Symp Quant Biol* 47:723–729
- Bateman AJ (1975) Simplification of palindromic telomere theory. *Nature* 253:379–380
- Borza T, Redmond EK, Laflamme M, Lee RW (2009) Mitochondrial DNA in the *Oogamochlamys* clade (Chlorophyceae): high GC content and unique genome architecture for green algae. *J Phycol* 45:1323–1334
- Boudreau E, Turmel M (1995) Gene rearrangements in *Chlamydomonas* chloroplast DNAs are accounted for by inversions and by the expansion/contraction of the inverted repeat. *Plant Mol Biol* 27:351–364
- Boudreau E, Otis C, Turmel M (1994) Conserved gene clusters in the highly rearranged chloroplast genomes of *Chlamydomonas moewusii* and *Chlamydomonas reinhardtii*. *Plant Mol Biol* 24:585–602
- 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
- Cavalier-Smith T (1974) Palindromic base sequences and replication of eukaryote chromosome ends. *Nature* 250:467–470
- Coleman AW, Thompson WF, Coff LJ (1991) Identification of the mitochondrial genome in the chrysophyte alga *Ochromonas danica*. *J Protozool* 38:129–135
- Denovan-Wright EM, Nedelcu AM, Lee RW (1998) Complete sequence of the mitochondrial DNA of *Chlamydomonas eugametos*. *Plant Mol Biol* 36:285–295
- Dinouël N, Drissi R, Miyakawa I, Sor F, Rousset S, Fukuhara H (1993) Linear mitochondrial DNAs of yeasts: closed-loop structure of the termini and possible linear-circular conversion mechanisms. *Mol Cell Biol* 13:2315–2323
- Fan J, Lee RW (2002) Mitochondrial genome of the colorless green alga *Polytomella parva*: two linear DNA molecules with homologous inverted repeat termini. *Mol Biol Evol* 19:999–1007
- Förstemann K, Höss M, Lingner J (2000) Telomerase-dependent repeat divergence at the 3' ends of yeast telomeres. *Nucleic Acids Res* 28:2690–2694
- González A, Talavera A, Almendral JM, Viñuela E (1986) Hairpin loop structure of African swine fever virus DNA. *Nucleic Acids Res* 14:6835–6844
- Gray MW, Boer PH (1988) Organization and expression of algal (*Chlamydomonas reinhardtii*) mitochondrial DNA. *Philos Trans R Soc B* 319:135–147
- Handa H (2007) Linear plasmids in plant mitochondria: peaceful coexistences or malicious invasions? *Mitochondrion* 8:15–25
- Hinnebusch J, Barbour AG (1991) Linear plasmids of *Borrelia burgdorferi* have a telomeric structure and sequence similar to those of a eukaryotic virus. *J Bacteriol* 173:7233–7239
- Kairo A, Fairlamb AH, Gobright E, Nene V (1994) A 7.1 kb linear DNA molecule of *Theileria parva* has scrambled rDNA sequences and open reading frames for mitochondrially encoded proteins. *EMBO J* 13:898–905
- Katz LA, Curtis EA, Pfunder M, Landweber LF (2000) Characterization of novel sequences from distantly related taxa by walking PCR. *Mol Phylogenet Evol* 14:318–321
- Kayal E, Lavrov DV (2008) The mitochondrial genome of *Hydra oligactis* (Cnidaria, Hydrozoa) sheds new light on animal mtDNA evolution and cnidarian phylogeny. *Gene* 410:177–186
- Kroymann J, Zetsche K (1998) The mitochondrial genome of *Chlorogonium elongatum* inferred from the complete sequence. *J Mol Evol* 47:431–440
- Lafflamme M, Lee RW (2003) Mitochondrial genome conformation among CW-group chlorophyte algae. *J Phycol* 39:213–220
- Lewis LA, McCourt M (2004) Green algae and the origin of land plants. *Am J Bot* 91:1535–1556
- Mallet MA, Lee RW (2006) Identification of three distinct *Polytomella* lineages based on mitochondrial DNA features. *J Eukaryot Microbiol* 53:79–84
- 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
- Miyashita S, Hirochika H, Ikeda JE, Hashiba T (1990) Linear plasmid DNAs of the plant pathogenic fungus *Rhizoctonia solani* with unique terminal structures. *Mol Gen Genet* 220:165–171
- Moore LJ, Coleman AW (1989) The linear 20 kb mitochondrial genome of *Pandorina morum* (Volvocaceae, Chlorophyta). *Plant Mol Biol* 13:459–465
- Nakada T, Misawa K, Nozaki H (2008) Molecular systematics of Volvocales (Chlorophyceae, Chlorophyta) based on exhaustive 18S rRNA phylogenetic analyses. *Mol Phylogenet Evol* 48:281–291
- Nash EA, Barbrook AC, Edwards-Stuart RK, Bernhardt K, Howe CJ, Nisbet RE (2007) Organization of the mitochondrial genome in the dinoflagellate *Amphidinium carterae*. *Mol Biol Evol* 24:1528–1536
- Nedelcu AM, Lee RW, Lemieux C, Gray MW, Burger G (2000) The complete mitochondrial DNA sequence of *Scenedesmus obliquus* reflects an intermediate stage in the evolution of the green algal mitochondrial genome. *Genome Res* 10:819–831
- Nosek J, Tomáška L (2002) Mitochondrial telomeres: alternative solutions to the end-replication problem. In: Krupp G, Parwarresch R (eds) *Telomerases telomeres and cancer*. Kluwer/Plenum Publishers, New York, pp 396–417
- Nosek J, Tomáška L (2003) Mitochondrial genome diversity: evolution of the molecular architecture and replication strategy. *Curr Genet* 44:73–84
- Nosek J, Tomáška L, Kucejová B (2004a) The chromosome end replication: lessons from mitochondrial genetics. *J Appl Biomed* 2:71–79
- Nosek J, Novotná M, Hlavatovicová Z, Ussery DW, Fajkus J, Tomáška L (2004b) Complete DNA sequence of the linear mitochondrial genome of the pathogenic yeast *Candida parapsilosis*. *Mol Genet Genom* 272:173–180
- Olovnikov AM (1971) Principle of marginotomy in template synthesis of polynucleotides. *Dokl Akad Nauk SSSR* 201:1496–1499
- Popescu CE, Lee RW (2007) Mitochondrial genome sequence evolution in *Chlamydomonas*. *Genetics* 175:819–826
- Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14:817–818
- Pringsheim EG (1955) The Genus *Polytomella*. *J Protozool* 2:137–145
- Pringsheim EG (1963) Farblose Algen. Ein Beitrag zur Evolutionsforschung. Gustav Fischer, Stuttgart
- Pritchard AE, Cummings DJ (1981) Replication of linear mitochondrial DNA from *Paramecium*: sequence and structure of the initiation-end crosslink. *Proc Natl Acad Sci USA* 78:7341–7345
- Pröschold T, Marin B, Schlösser UG, Melkonian M (2001) Molecular phylogeny and taxonomic revision of *Chlamydomonas* (Chlorophyta). I. Emendation of *Chlamydomonas* Ehrenberg and *Chloromonas* Gobi, and description of *Oogamochlamys* gen. nov. and *Lobochlamys* gen. nov. *Protist* 152:265–300
- Rohozinski J, Girton LE, Van Etten JL (1989) *Chlorella* viruses contain linear nonpermuted double-stranded DNA genomes with covalently closed hairpin ends. *Virology* 168:363–369
- Ryan R, Grant D, Chiang KS, Swift H (1978) Isolation and characterization of mitochondrial DNA from *Chlamydomonas reinhardtii*. *Proc Natl Acad Sci USA* 75:3268–3272
- Shao Z, Graf S, Chaga OY, Lavrov DV (2006) Mitochondrial genome of the moon jelly *Aurelia aurita* (Cnidaria, Scyphozoa): a linear DNA molecule encoding a putative DNA-dependent DNA polymerase. *Gene* 381:92–101
- Shukla GC, Nene V (1998) Telomeric features of *Theileria parva* mitochondrial DNA derived from cycle sequence data of total genomic DNA. *Mol Biochem Parasitol* 95:159–163
- Smith DR, Lee RW (2008a) Nucleotide diversity in the mitochondrial and nuclear compartments of *Chlamydomonas reinhardtii*: investigating the origins of genome architecture. *BMC Evol Biol* 8:156
- Smith DR, Lee RW (2008b) Mitochondrial genome of the colorless green alga *Polytomella capuana*: a linear molecule with an unprecedented GC content. *Mol Biol Evol* 25:487–496
- Smith DR, Lee RW (2009) The mitochondrial and plastid genomes of *Volvox carteri*: bloated molecules rich in repetitive DNA. *BMC Genom* 10:132
- Smith DR, Lee RW (2010) Low nucleotide diversity for the expanded organelle and nuclear genomes of *Volvox carteri* supports the mutational-hazard hypothesis. *Mol Biol Evol* 27:1–13

- Smith DR, Lee RW, Cushman JC, Magnuson JK, Tran D, Polle JEW (2010) The *Dunaliella salina* organelle genomes: large sequences, inflated with intronic and intergenic DNA. *BMC Plant Biol* 10:83
- Steinhauser S, Beckert S, Capesius I, Malek O, Knoop V (1999) Plant mitochondrial RNA editing. *J Mol Evol* 48:303–312
- Swofford DL (2003) PAUP\*: phylogenetic analysis using parsimony (and other methods, 4.0 Beta. Sinauer Associates, Sunderland
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680
- Traktman P (1996) Poxvirus DNA replication. In: DePamphilis ML (ed) DNA replication in eukaryotic cells. Cold Spring Harbour Laboratory Press, Cold Spring Harbor, pp 775–798
- Turmel M, Otis C, Lemieux C (2010) A deviant genetic code in the reduced mitochondrial genome of the picoplanktonic green alga *Pycnococcus provasolii*. *J Mol Evol* 70:203–214
- Vahrenholz C, Riemen G, Pratje E, Dujon B, Michaelis G (1993) Mitochondrial DNA of *Chlamydomonas reinhardtii*: the structure of the ends of the linear 15.8-kb genome suggests mechanisms for DNA replication. *Curr Genet* 24:241–247
- Voigt O, Erpenbeck D, Wörheide G (2008) A fragmented metazoan organellar genome: the two mitochondrial chromosomes of *Hydra magnipapillata*. *BMC Genom* 9:350
- Watson JD (1972) Origin of concatemeric T7 DNA. *Nat New Biol* 239:197–201