DOI: 10.1534/genetics.106.063156

Mitochondrial Genome Sequence Evolution in Chlamydomonas

Cristina E. Popescu and Robert W. Lee[1]

Department of Biology, Dalhousie University, Halifax, Nova Scotia B3H 4J1, Canada

Manuscript received July 8, 2006
Accepted for publication November 15, 2006

ABSTRACT

The mitochondrial genomes of the Chlorophyta exhibit significant diversity with respect to gene content and genome compactness; however, quantitative data on the rates of nucleotide substitution in mitochondrial DNA, which might help explain the origin of this diversity, are lacking. To gain insight into the evolutionary forces responsible for mitochondrial genome diversification, we sequenced to near completion the mitochondrial genome of the chlorophyte *Chlamydomonas incerta*, estimated the evolutionary divergence between *Chlamydomonas reinhardtii* and *C. incerta* mitochondrial protein-coding genes and rRNA-coding regions, and compared the relative evolutionary rates in mitochondrial and nuclear genes. Synonymous and nonsynonymous substitution rates do not differ significantly between the mitochondrial and nuclear protein-coding genes. The mitochondrial rRNA-coding regions, however, are evolving much faster than their nuclear counterparts, and this difference might be explained by relaxed functional constraints on the mitochondrial translational apparatus due to the small number of proteins synthesized in Chlamydomonas mitochondria. Substitution rates at synonymous sites in a nonstandard mitochondrial gene (*rtl*) and at intronic and synonymous sites in nuclear genes expressed at low levels suggest that the mutation rate is similar in these two genetic compartments. Potential evolutionary forces shaping mitochondrial genome evolution in Chlamydomonas are discussed.

KNOWLEDGE of the rates of synonymous and nonsynoymous substitutions in mitochondrial genes, both in absolute terms and relative to nuclear genes, has been used to make hypotheses about the evolutionary forces giving rise to the extensive diversity in coding capacity (Berg and Kurland 2000; reviewed by Adams and Palmer 2003) and the amount of intronic and intergenic DNA (Lynch *et al.* 2006) among mitochondrial genomes from different lineages. Central among these forces is the mutation rate, commonly estimated from the substitution rate at synonymous sites, assuming that these sites are evolving neutrally (Kimura 1983). Moreover, due to the effectively asexual nature of the mitochondrial genome, it has been suggested that the rate of accumulation of mildly deleterious mutations is higher in the mitochondrial compartment compared to the nuclear one and it has been proposed that this phenomenon, especially under conditions of high mutation rate, could contribute further to a low coding capacity of mitochondrial genomes by favoring the functional transfer of mitochondrial genes to the nucleus (*e.g.*, Lynch 1996; Andersson and Kurland 1998; Lynch and Blanchard 1998; Martin and Herrmann 1998). A higher rate of accumulation of mildly deleterious mutations in the mitochondrial *vs.* nuclear genes should be reflected in an increased rate of nucleotide substitutions relative to the rate of mutation in the former than in the latter compartment or, more specifically, in a higher ratio of nonsynonymous per synonymous substitutions in the mitochondrial than in the nuclear protein-coding genes (Lynch and Blanchard 1998).

Mitochondrial genome sequence data from the Chlorophyta, which includes most of the green algal diversification in the Plantae (Keeling *et al.* 2005), *i.e.*, the chlorophyceans, trebouxiophyceans, ulvophyceans, and prasinophyceans (Lewis and McCourt 2004), have revealed a fivefold variation in gene content and a genome architecture that varies from compact to expanded (Bullerwell and Gray 2004; Gray *et al.* 2004; Pombert *et al.* 2004, 2006). Within this group, the reported number of standard protein-coding genes in mitochondrial DNA (mtDNA) ranges from a low of seven, such as in *Chlamydomonas reinhardtii* and other closely related chlorophycean taxa, to a high of 34 as seen in the prasinophycean *Nephroselmis olivacea*, and the average proportion of intronic plus intergenic DNA varies from ~10% in *C. reinhardtii* (not counting the inverted terminal repeats) to >50% for the ulvophycean *Pseudendoclonium alkinetum* (Pombert *et al.* 2004, 2006; Lynch *et al.* 2006). Members within the Chlorophyta differ from those within most eukaryotic phyla, which have rather uniform mitochondrial gene contents (Adams and Palmer 2003; Bullerwell and Gray 2004). For all these reasons, chlorophytes appear to be useful for testing hypotheses that lineage-specific rates of mutation and the rate of accumulation of deleterious

[1]*Corresponding author:* Department of Biology, Dalhousie University, Halifax, NS B3H 4J1, Canada. E-mail: robert.lee@dal.ca

mutations in the mitochondrial and nuclear genomes have determined the degree of shrinkage in coding capacity and compactness of the mitochondrial genome.

Although there has been no quantitative study of the rates of nucleotide substitution in mitochondrial genes in the Chlorophyta, the mtDNA of *C. reinhardtii* is thought to be evolving more rapidly than the nuclear DNA (nDNA), similar to the situation described in mammals (Brown *et al.* 1982; Pesole *et al.* 1999) but opposite what is typically understood for land plants (Wolfe *et al.* 1987; Gaut 1998; Muse 2000). This belief is based on phylogenic analyses that support in *C. reinhardtii* an accelerated rate of substitution in the small subunit (SSU) and the large subunit (LSU) ribosomal RNA (rRNA)-coding regions in the mitochondrial compartment relative to those in the nucleus (Gray *et al.* 1989).

The mitochondrial genome of *C. reinhardtii* is a 15,758-bp linear DNA molecule containing no introns, a low fraction of intergenic DNA, 12 standard genes, and one nonstandard gene (Gray and Boer 1988; Michaelis *et al.* 1990). The standard genes encode seven respiratory chain proteins, three transfer RNAs (tRNAs), and the SSU and LSU rRNAs. The rRNA genes are broken into several pieces (four for the SSU and eight for the LSU) interspersed with one another and with protein- and tRNA-coding regions (Boer and Gray 1988a). The nonstandard gene in *C. reinhardtii* mtDNA has been described as a reverse transcriptase-like (*rtl*) coding sequence (Boer and Gray 1988b) whose function and evolutionary origin remain uncertain.

In this study, we have undertaken the sequencing of the *Chlamydomonas incerta* mtDNA to estimate rates of nucleotide substitution between *C. reinhardtii* and *C. incerta* mitochondrial protein-coding genes and to compare these rates with those recently reported for the nuclear genes from the same taxa (Popescu *et al.* 2006). This work also measures the rates of substitution in the mitochondrial- and nuclear-encoded SSU and LSU rRNA-coding regions between the two taxa.

MATERIALS AND METHODS

***C. incerta* strain and culture conditions:** *C. incerta* was obtained from the Sammlung von Algenkulturen, Göttingen (SAG), Germany, where it is listed as SAG 7.73 under the name *C. reinhardtii* on the basis of morphological criteria and susceptibility to autolysin from the *C. reinhardtii* group (SAG, personal communication). Cells were cultured in liquid minimal medium under alternating light/dark cycles and total cellular DNA was prepared as described by Laflamme and Lee (2003).

Generation of DNA sequences: A PCR-based approach was employed to sequence an internal 15,637-bp fragment of the *C. incerta* mitochondrial genome. Platinum PCR SuperMix High Fidelity kit (Invitrogen, San Diego) was used to amplify initially seven mtDNA segments >2 kbp. The sets of primers were designed primarily from *C. reinhardtii* mtDNA (GenBank accession no. NC_001638). Annealing temperatures were from 55° to 60°. PCR products were gel purified and recovered with a gel extraction kit (QIAGEN, Chatsworth, CA). Some PCR products were cloned and sequenced, while others were sequenced directly (Center for Applied Genomics, Hospital for Sick Children, Toronto). All sequences reported correspond to at least two independent clones or PCR products. Additional PCR reactions were run and the products were sequenced directly to resolve any ambiguities or to fill gaps in the sequence. The annotation of regions in *C. incerta* mtDNA is based on comparisons with homologs in *C. reinhardtii* mtDNA. The nuclear SSU rRNA gene of *C. incerta* was obtained from GenBank (AY781664) and the sequence was confirmed by sequencing independent PCR products. The portion of the *C. incerta* nuclear LSU rRNA gene encoding 28S rRNA was obtained by sequencing both strands of several overlapping PCR products. Moreover, authentication of the *C. incerta* rDNA sequences was based on comparisons with the EST sequences retrieved from our *C. incerta* cDNA libraries (Popescu *et al.* 2006) that correspond to the nuclear SSU and LSU rRNA-coding regions. The nuclear SSU and LSU (28S) rRNA-coding regions of *C. reinhardtii* were accessed from GenBank (M32703) and ChlamyDB at http://www.chlamy.org/chlamydb.html, respectively.

Data analysis: CODONS (Lloyd and Sharp 1992), MEGA 3.1 (Kumar *et al.* 2004), and DAMBE (Xia and Xie 2001) software packages were used to compute the effective number of codons (Wright 1990), base composition, and relative synonymous codon usage, respectively. *C. reinhardtii* and *C. incerta* homologous mtDNA sequences were aligned using ClustalX (Thompson *et al.* 1997). The number of synonymous substitutions per synonymous site and the number of nonsynonymous substitutions per nonsynonymous site in the mitochondrial protein-coding regions were estimated using the maximum-likelihood method (Goldman and Yang 1994) implemented in the CODEML program of the version 3.14 PAML package (Yang 1997); the model used accounts for transition/transversion bias and estimates the expected codon frequencies using the nucleotide frequencies at each codon position (F3x4). Distance estimates for the protein-coding nuclear genes under the same codon substitution model were taken from (Popescu *et al.* 2006). In a separate set of analyses of both mitochondrial and nuclear protein-coding genes, codon frequencies were used as free parameters (F61 model); the results led to the same conclusions and therefore are not presented. The number of substitutions per site in the mitochondrial and nuclear rRNA genes was calculated using the Hasegawa–Kishino–Yano (HKY85) model (Hasegawa *et al.* 1985) implemented in the BASEML program, which is also part of the PAML package. The structural cores of the *C. reinhardtii* and *C. incerta* mitochondrial- and nuclear-encoded rRNAs were defined on the basis of the highly conserved blocks of the rRNA sequences as characterized at the comparative RNA website (http://www.rna.icmb.utexas.edu/) (Cannone *et al.* 2002). The secondary structure of the core regions of the *C. reinhardtii* and *C. incerta* mitochondrial- and nuclear-encoded rRNAs and the alignments of the homologous rRNA genes are presented as supplemental Figures S1–S10 at http://www.genetics.org/supplemental/. The paired *t*-tests and the calculation of Pearson correlation coefficients were performed using MINITAB, release 14.12.0.

Nucleotide sequence accession numbers: The partial sequence of the *C. incerta* mitochondrial genome and the partial sequence of the *C. incerta* nuclear-encoded LSU rRNA gene have been deposited in GenBank under accession nos. DQ373068 and DQ373067, respectively.

RESULTS

Mitochondrial genome structure: An internal 15,760-bp segment of the *C. incerta* mtDNA was sequenced.

This region is estimated to represent ~87% of the complete size of the *C. incerta* mtDNA, predicted by Southern blot analyses to be ~18,000 bp (our unpublished data). All coding regions previously identified in *C. reinhardii* mtDNA were identified in the partially sequenced mtDNA of *C. incerta* (supplemental Figure S11 at http://www.genetics.org/supplemental/). In *C. incerta*, these coding regions, as in the *C. reinhardtii* counterpart, are densely packed. Relative to the *C. reinhardtii* mtDNA coordinates, the partially sequenced *C. incerta* mtDNA does not include the left terminal repeat sequence and 438 bp of the 3′-end of *cob* and the right terminal repeat sequence and 113 bp of the 3′-end of *rrnL2b*. The *C. incerta* mtDNA is colinear with its *C. reinhardtii* homolog, and its ~2000 bp larger size can be explained by two group I introns, which are present in the mitochondrial genome of *C. incerta* but are missing from the *C. reinhardtii* homolog; one of the introns is in the gene for apocytochrome b (*cob*) and the other in the gene for subunit 1 of cytochrome c oxidase (*cox1*). Each intron sequence contains an open reading frame, which shows similarity at the predicted amino acid level to the LAGLIDADG endonuclease family. Another difference between the mtDNA of the two species concerns the coding module for the putative LSU rRNA fragment L3a (*rrnL3a*). In *C. reinhardtii*, *rrnL3a* is a 114-bp sequence, which contains an imperfect inverted repeat sequence (Boer and Gray 1991) and is immediately upstream of and cotranscribed with *rrnL3b*, a region encoding a core fragment of the LSU rRNA; the L3a/L3b cotranscript is separated by post-transcriptional processing (Boer and Gray 1988a). In *C. incerta*, the *rrnL3a* coding region, also located immediately upstream of *rrnL3b* (supplemental Figure S11 at http://www.genetics.org/supplemental/), is only 65 bp long and flanked at its 5′-end by an A-rich, 11-nucleotide sequence that can form a stable hairpin structure with a sequence near the 3′-end of this coding region.

Mitochondrial codon usage: *C. incerta* mitochondrial genes like those of *C. reinhardtii* use the universal genetic code, and synonymous codon usage of mitochondrial genes does not differ significantly between these two taxa. The effective number of codons (N_c) varies only slightly among mitochondrial genes in the two species and these values are positively and strongly correlated among orthologous gene pairs ($r = 0.97$). The mean N_c values across all mitochondrial genes are not significantly different between the two species ($t = 1.16$, $P = 0.28$) (Table 1). In *C. incerta*, as reported earlier for *C. reinhardtii*, the standard mitochondrial protein-coding genes exhibit strong synonymous codon usage bias (low N_c). In the standard mitochondrial protein-coding genes, five of eight fourfold degenerate codon families favor T-ending followed by C-ending codons, while three families favor A-ending followed by T- or G-ending ones; *rtl* uses the synonymous codons more randomly than the standard genes (supplemental Table S1 at http://www.genetics.org/supplemental/). In terms of N_c and relative synonymous codon usage, we found no obvious difference between the group of standard mitochondrial genes encoded on one strand (*cob*, *nad4*, and *nad5*) and the group encoded on the other strand (*cox1*, *nad2*, *nad6*, and *nad1*).

TABLE 1

The effective number of codons (N_c) in mitochondrial protein-coding genes

	N_c	
Genes	*C. incerta*	*C. reinhardtii*
Standard		
cob[a]	31.05	32.15
cox1	31.34	30.29
nad1	36.03	36.49
nad2	39.87	38.09
nad4	36.65	34.22
nad5	35.88	36.48
nad6	33.78	31.74
Average[b]	34.94 (3.13)	34.21 (2.92)
Nonstandard		
rtl	46.68	47.21

[a] Partial sequence.
[b] The average is the arithmetic mean and values in parentheses are the standard deviations computed over all genes.

The base composition at fourfold degenerate sites in protein-coding genes was compared to the base composition of introns and intergenic regions. Among all protein-coding genes, T and A, in that order, are the preferred nucleotides at fourfold sites and this bias is greater in the standard protein-coding genes than in *rtl*; the bias favoring T and A is less evident in the intronic and intergenic regions (Table 2).

Rates of nucleotide substitution in mitochondrial genes: The relative nucleotide substitution rates for the different mitochondrial genes are reported in Tables 3 and 4. Among the standard mitochondrial protein-coding genes, estimates of the number of synonymous substitutions per synonymous site (d_S) vary only slightly, whereas estimates of the number of nonsynonymous substitutions per nonsynonymous site (d_N) show 30-fold variation with *nad2* and *nad4* having the lowest and highest values, respectively. The nonstandard coding region, *rtl*, evolves more rapidly than the fastest-evolving standard mitochondrial gene in terms of both d_S and d_N. No correlation has been found between the synonymous and nonsynonymous substitution rates of *C. reinhardtii* and *C. incerta* mitochondrial protein-coding sequences ($r = 0.36$, $P = 0.39$). The d_N/d_S ratios for the individual standard mitochondrial genes as well as for the nonstandard *rtl* are all considerably <1, but among these, *rtl* has the highest ratio. The nucleotide substitution rates for the mitochondrial SSU and LSU rRNA genes were estimated separately for the conserved core

TABLE 2

Base composition (%) at fourfold degenerate sites in protein-coding genes and in noncoding regions of the mitochondrial genome

Species	Protein-coding genes (no. of sites)	A	T	G	C
C. incerta	Standard[a] (1222)	27.5	52.9	3.7	15.9
	rtl (155)	25.2	42.6	14.2	18.1
C. reinhardtii	Standard[a] (1290)	26.3	52.3	2.2	19.1
	rtl (160)	24.4	40	15	20.6
Species	Noncoding regions (no. of sites)	A	T	G	C
C. incerta	Intron *cob* (1078)	28.9	29.9	20.3	20.9
	Intron *cox1* (1035)	27.2	29.2	23.8	19.8
C. reinhardtii	Intergenic[b] (916)	29.1	29.7	18.1	23.1
	Intergenic[b] (1214)	28.6	30.4	19.9	21.1

[a] Concatenated sequences.
[b] Concatenated sequences not including the region between *nad5* and *cox1* where a switch in the template strands occurs.

and the variable regions (Table 4). The conserved regions of each gene evolve at less than half the rate of the variable regions in the same gene, and estimates of the number of substitutions per site for both the conserved core and variable regions are slightly higher for the LSU gene compared to the SSU gene. Overall substitution rates in the SSU and LSU mitochondrial rRNA-coding regions are approximately three to seven times higher than the rate at nonsynonymous sites averaged over all standard mitochondrial protein-coding genes. Even the conserved core sequences of the SSU and LSU rRNAs evolve faster than the nonsynonymous sites in all but the fastest-evolving standard mitochondrial protein-coding gene, *nad4*.

TABLE 3

Nucleotide substitution rates (± standard errors) between *C. reinhardtii* and *C. incerta* mitochondrial protein-coding genes

Genes	d_S	d_N	d_N/d_S
Standard			
cob[a]	0.22 ± 0.04	0.004 ± 0.003	0.018
cox1	0.32 ± 0.04	0.011 ± 0.003	0.034
nad1	0.25 ± 0.04	0.002 ± 0.002	0.006
nad2	0.31 ± 0.04	0.001 ± 0.001	0.004
nad4	0.26 ± 0.03	0.036 ± 0.007	0.137
nad5	0.38 ± 0.04	0.012 ± 0.003	0.032
nad6	0.28 ± 0.06	0.012 ± 0.006	0.044
Average[b]	0.29 (0.05)	0.011 (0.011)	0.039 (0.046)
Nonstandard			
rtl	0.6 ± 0.07	0.17 ± 0.017	0.28

d_S, number of synonymous substitutions/synonymous site. d_N, number of nonsynonymous substitutions/nonsynonymous site.
[a] Partial sequence.
[b] The average is the arithmetic mean and values in parentheses are the standard deviations computed over all genes.

Rates of nucleotide substitution in nuclear genes: For the nuclear compartment, data on the sequence divergence between *C. reinhardtii* and *C. incerta* in 67 protein-coding gene sequences were taken from Popescu *et al.* (2006), while data on the sequence divergence in SSU and LSU rRNA-coding regions are estimated here (Table 4). The nuclear rRNA-coding regions of *C. reinhardtii* are almost identical with their *C. incerta* counterparts; no substitution differences were found between the corresponding conserved core regions, and changes of only four and one nucleotide were noted in the variable regions of the SSU and LSU rDNAs, respectively. The overall divergence of the nuclear rRNA genes in the *C. reinhardtii* and *C. incerta* comparison is equivalent to that of the most slowly evolving nuclear protein-coding genes in terms of d_N.

Rates of nucleotide substitution in mitochondrial *vs.* nuclear genes: Estimates of the relative nucleotide substitution rates in protein-coding genes and rRNA-coding regions from the mitochondrial and nuclear genetic compartments are summarized in Table 4. On average, there is no statistically significant difference between the standard mitochondrial- and nuclear-encoded protein-coding genes in terms of synonymous ($t = 0.73$, $P = 0.47$) and nonsynonymous ($t = 0.83$, $P = 0.41$) rates. Similar results were obtained when *rtl* was included in the analysis (data not shown). Nevertheless, because synonymous substitution rates vary widely among the nuclear genes, unlike the synonymous rates among the mitochondrial genes, and because nonsynonymous rates are very heterogeneous among genes in both compartments, the ratio of the mitochondrial and nuclear evolutionary rates in protein-coding genes varies considerably with the particular genes compared (Table 4). Significantly higher nucleotide substitution rates were found in the mitochondrial SSU and LSU rRNA-coding regions compared to the nuclear counterparts, and this is true whether conserved or variable regions are considered (Table 4).

DISCUSSION

Comparative genome organization of *C.reinhardtii* and *C. incerta* mtDNA: On the basis of the large segment of the *C. incerta* mtDNA sequenced in this study, the *C. reinhardtii* and *C. incerta* mitochondrial genomes have a similar gene content and genomic organization. The only notable differences identified are the presence of two group I introns and the absence of the 5′ half of the putative coding module *rrnL3a* in the *C. incerta* mtDNA compared to the *C. reinhardtii* counterpart.

TABLE 4

Comparison of nucleotide substitution divergence (± standard errors) between *C. reinhardtii* and *C. incerta* mitochondrial and nuclear genes

Type of site	mtDNA	nDNA	Ratio of mtDNA/nDNA
Synonymous[a]			
Average[b]	0.29 (0.05)	0.37 (0.29)	0.78
Minimum	0.22 ± 0.04	0.025 ± 0.018	6.66
Maximum	0.38 ± 0.04	1.68 ± 0.48	0.23
Nonsynonymous[a]			
Average[b]	0.011 (0.01)	0.018 (0.02)	0.61
Minimum	0.001 ± 0.001	0	—
Maximum	0.036 ± 0.007	0.120 ± 0.015	0.30
Nonsynonymous/synonymous[a]			
Average[b]	0.039 (0.046)	0.056 (0.04)	0.70
Minimum	0.004	—	—
Maximum	0.137	0.217	0.63
SSU rRNA			
Overall	0.036 ± 0.007	0.002 ± 0.001	18
Conserved core	0.021 ± 0.005	0	—
Variable regions	0.076 ± 0.017	0.0044 ± 0.002	17
LSU rRNA			
Overall	0.075 ± 0.008	0.0004 ± 0.0003	188
Conserved core	0.040 ± 0.004	0	—
Variable regions	0.120 ± 0.008	0.0006 ± 0.0006	200

[a] Standard mitochondrial and nuclear protein-coding genes.
[b] The average is the arithmetic mean and the values in parentheses are the standard deviations computed over all genes.

The presence of the group I intron encoding the LAGLIDADG motif in the *cob* gene of both *C. incerta* (this study) and an interfertile relative of *C. reinhardtii*, namely, *Chlamydomonas smithii* (Colleaux *et al.* 1990; Ma *et al.* 1992), suggests that the last ancestor of the three taxa harbored such an intron and that *C. reinhardtii cob* lost the intron. The presence in *C. incerta* mtDNA of a truncated *rrnL3a* sequence may not be surprising as no *rrnL3a* homolog has been identified in the mtDNA of other more distantly related chlorophyceans (Denovan-Wright and Lee 1994; Fan *et al.* 2003). Actually, at the time of its description, it was considered questionable whether or not the L3a transcript functions in the mitochondrial ribosome of *C. reinhardtii* (Boer and Gray 1988a) as this sequence showed no recognizable counterpart in the LSU rRNA of *Escherichia coli* or any other taxon. Boer and Gray (1991) proposed that imperfect inverted repeat sequences identified in the mtDNA of *C. reinhardtii* may play a role in the processing of mitochondrial transcripts. Both *rrnL3a* in *C. reinhardtii* (Boer and Gray 1991) and its truncated homolog in *C. incerta*, as determined here, contain such repeat sequences; therefore we suggest that the *rrnL3a* segments are retained for RNA processing and do not encode constituents of the mitochondrial ribosomes.

Mitochondrial codon usage: The strong codon usage bias of the standard mitochondrial genes of *C. reinhardtii* and *C. incerta* might be best understood in terms of the selection–mutation–drift theory (reviewed by Rocha 2004), which attempts to explain the coevolution of codon usage bias and tRNA abundance in the context of translational optimization (but also see Xia 2005). In *C. incerta*, as in *C. reinhardtii*, only three mitochondrial-encoded tRNAs, namely, $tRNA^{Met}$, $tRNA^{Trp}$, and $tRNA^{Gln}$, were identified; a total of at least 22 or 23 tRNAs must be imported into the mitochondria, depending on whether single or separate initiator and elongator $tRNAs^{Met}$ are used (Boer and Gray 1988c). The mitochondrial protein-coding genes in both species employ preferentially T- or A-ending codons in fourfold degenerate families. This could be explained in part by strand-specific mutation pressure favoring T and A, which is evident in the intergenic and intron regions of the mtDNA. On the other hand, there may be selection for nucleotides in the wobble position of the anticodons to provide maximum "flexibility" in anticodon–codon pairing and thereby limit the number of tRNAs needed as suggested for vertebrate mitochondria (Xia 2005). The analysis of *rtl* codon usage in *C. incerta*, as shown earlier for *C. reinhardtii* (Boer and Gray 1988b,c), revealed less codon bias than found in the standard mitochondrial protein-coding genes and the presence of several unique codons that are not used by any of the standard genes (two of these unique codons are shared between *C. incerta* and *C. reinhardtii rtl*). These data suggest that codon usage in *rtl* is under little translational selection in contrast to the standard protein-coding genes.

Rates of nucleotide substitution in mitochondrial protein-coding genes: The standard mitochondrial protein-coding genes of the Chlamydomonas taxa examined show little variation in the synonymous substitution rates but large variation in the nonsynonymous substitution rates. The lack of variation in synonymous rates among genes suggests that all mitochondrial standard protein-coding genes experience a similar mutation rate and have comparable selective constrains at synonymous sites; the differences in the nonsynonymous rates are likely due to differences in the strength of purifying selection as a result of variation in functional constraints. Low synonymous substitution rate variation among mitochondrial genes and high variation in nonsynonymous rates were also reported in mammals (Pesole *et al.* 1999) and land plants (Laroche *et al.* 1997).

Our analyses uncovered the *rtl* locus as more rapidly evolving than standard mitochondrial protein-coding genes at both synonymous and nonsynonymous sites. The observations that *rtl* has the highest d_S value and little codon usage bias are consistent with the hypothesis that synonymous sites in this gene are under very weak selective constraints. The higher d_N/d_S ratio for *rtl* compared to those for standard mitochondrial protein-coding genes suggests that nonsynonymous sites in this gene are also under less functional selective constraints compared to the standard mitochondrial protein-coding genes, so that nonsynonymous changes in *rtl* are not selected against and therefore accumulate. Interestingly, *rtl* has not been identified in other completely sequenced (Denovan-Wright *et al.* 1998; Kroymann and Zetsche 1998; Kück *et al.* 2000; Nedelcu *et al.* 2000) and nearly completely sequenced (Fan and Lee 2002) chlorophycean mitochondrial genomes. Nevertheless, there is evidence that *rtl* is transcribed in *C. reinhardtii* (Boer and Gray 1988b), and as observed in this study, there are no frameshifting changes in the *C. reinhardtii* and *C. incerta rtl* sequences, in spite of the presence of seven additional codons at two different locations in the 5′-end of the *C. incerta* sequence relative to the *C. reinhardtii* homolog and of numerous point substitution differences. These results suggest that *rtl* is not a pseudogene. It has been proposed that *rtl* in *C. reinhardtii* may encode a reverse-transcriptase-like protein essential for the maintenance of the ends of the linear mtDNA (Vahrenholz *et al.* 1993) and that this gene is a remnant of a group II intron-encoded open reading frame (Nedelcu and Lee 1998). In the future, questions about the role, if any, of *rtl* in these green algal mtDNAs might be answered by using reverse genetics if procedures of mitochondrial transformation using *C. reinhardtii* continue to improve (Yamasaki *et al.* 2005; Remacle *et al.* 2006).

Rates of nucleotide substitution in mitochondrial *vs.* nuclear genes: Available data on substitution rates in the mitochondrial and nuclear genomes of Chlamydomonas suggest that the mutation rates in the mitochondrial and nuclear genetic compartments are rather similar. The synonymous substitution rate in *rtl*, the rapidly evolving mitochondrial-coding gene with low codon usage bias, may be the best measure of the mutation rate in the mitochondrion, while substitution rates at intronic and synonymous sites in nuclear genes that are lowly expressed may give the best estimate of the mutation rate in the nucleus (Popescu *et al.* 2006). In this respect, Chlamydomonas differs from mammals (Brown *et al.* 1982; Pesole *et al.* 1999) and land plants (Wolfe *et al.* 1987; Gaut 1998; Muse 2000, but see also Palmer *et al.* 2000; Cho *et al.* 2004; Parkinson *et al.* 2005) where the estimated mutation rate is typically much higher and lower, respectively, in the mtDNA than in the nDNA. Nevertheless, Chlamydomonas resembles other unicellular eukaryotes outside the Plantae, which show little difference between the mutation rates in the mitochondrial and nuclear genetic compartments (Lynch *et al.* 2006).

Relative rates of mutation in the mitochondrial and nuclear genomes may affect the rate at which mitochondrial genes are functionally transferred to the nucleus. One hypothesis proposes that there is selection for the movement of genes from the mitochondria to the nucleus to escape the accumulation of deleterious mutations in the asexual mitochondrial genome (Muller's ratchet) (Muller 1964), especially in lineages with a much higher mutation rate in the mtDNA than in the nDNA (Lynch 1996; Adams and Palmer 2003). Another model (Berg and Kurland 2000) proposes that the functional transfer of mitochondrial genes to the nucleus is driven by cellular mechanisms that favor the transfer of mitochondrial gene copies to the nucleus rather than vice versa (see Thorsness and Fox 1990, 1993 for experimental evidence in yeast). According to this model, genes can be inactivated in the mitochondria and their copies fixed in the nucleus even when the rate of mutation is lower in the mitochondria than in the nucleus, but the rate of this partitioning could increase as the relative rate of mutation in the mtDNA increases. Because the relative mutation rates seem to be similar in the mitochondrial and nuclear genetic compartments of Chlamydomonas and the same appears to be true for the nonsynonymous rates, the low mtDNA-coding capacity cannot be explained by the first model; however, equal mutation rates in the two compartments might be consistent with the second model if future work shows that green algal lineages with fivefold greater content of mitochondrial genes experience lower rates of mutation in mtDNA than in nDNA.

The evolution of genome compactness, *i.e.*, the proportion of DNA associated with intronic and intergenic regions, has been hypothesized to be influenced by the power of random genetic drift and the mutation rate (Lynch *et al.* 2006; Lynch 2006). *C. reinhardtii* has a compact mitochondrial genome containing no introns

and ~10% intergenic DNA, not counting the terminal repeat sequences (Gray and Boer 1988; Michaelis *et al.* 1990), while the proportion of intron plus intergenic DNA in the nuclear genome is ~70% on the basis of linkage group III sequence (Li *et al.* 2003). This dissimilarity in genome compactness cannot be explained by mutation rate differences because the estimated mutation rates in the two genetic compartments seem to be quite similar. Other factors that might explain this difference include a possible reduction in the power of random genetic drift in the mitochondrial compartment compared to the nuclear one as discussed by Lynch *et al.* (2006) or selection for mitochondrial genome streamlining because of competition among mtDNA molecules for faster replication time (Selosse *et al.* 2001).

In contrast to the similarity in the average evolutionary rates of protein-coding genes between the two compartments, the evolutionary divergence of SSU and LSU rRNA genes in both the conserved and variable regions is much greater for the mitochondrial compartment compared to the nuclear one. Previous phylogenetic analyses revealed that the conserved core sequences of the mitochondrial rRNA genes of *C. reinhardtii* are more rapidly evolving than their nuclear-encoded counterparts (Gray *et al.* 1989). Nevertheless, as this study shows, the rRNA-coding regions are poor indicators of the overall relative trends of evolutionary change in the mitochondrial and nuclear compartments of Chlamydomonas. Our data suggest that the observed difference in the evolutionary rates of the rRNA-coding regions in the mitochondria *vs.* the nucleus in Chlamydomonas is largely the result of a difference in the strength of purifying selection acting on these coding regions in the two compartments. It was previously suggested that genes for components of the translational machinery in the mitochondrion compared to the cytosol could be freer to accumulate substitutions if the number of proteins synthesized in the mitochondrion was small, and if none or only a few of these proteins were involved in the mitochondrial information-transfer processes (Cann *et al.* 1984; Wilson *et al.* 1985). Indeed, the mitochondrial genome of *C. incerta* and *C. reinhardtii* contains only SSU and LSU rRNA-coding regions, seven standard protein-coding genes, and no protein-coding genes involved in information transfer. Moreover, the mitochondrial rRNA genes in these taxa are also highly fragmented, which might be another indication of relaxed functional constraints on the mitochondrial translational apparatus. The low evolutionary rates of the nuclear rRNA genes, as in other eukaryotic taxa, might best be explained by strong purifying selection rather than by concerted evolution of this multicopy gene family (Nei and Rooney 2005).

We thank Murray Schnare for a detailed editing of the SSU and LSU rRNA secondary structure models and for his many comments and suggestions used in the preparation of this article. We also thank Tudor Borza and Joseph Bielawski for helpful discussions during the course of this work and Ahalya Rajeswaran for editorial suggestions. C.E.P. was supported by scholarships from Dalhousie University and the Patrick Lett Fund. This work was funded by a grant from the Natural Sciences and Engineering Research Council of Canada to R.W.L.

LITERATURE CITED

Adams, K. L., and J. D. Palmer, 2003 Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. Mol. Phylogenet. Evol. **29:** 380–395.

Andersson, S. G., and C. G. Kurland, 1998 Reductive evolution of resident genomes. Trends Microbiol. **6:** 263–268.

Berg, O. G., and C. G. Kurland, 2000 Why mitochondrial genes are most often found in nuclei. Mol. Biol. Evol. **17:** 951–961.

Boer, P. H., and M. W. Gray, 1988a Scrambled ribosomal RNA gene pieces in *Chlamydomonas reinhardtii* mitochondrial DNA. Cell **55:** 399–411.

Boer, P. H., and M. W. Gray, 1988b 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* mitochondrial DNA. EMBO J. **7:** 3501–3508.

Boer, P. H., and M. W. Gray, 1988c Transfer RNA genes and the genetic code in *Chlamydomonas reinhardtii* mitochondria. Curr. Genet. **14:** 583–590.

Boer, P. H., and M. W. Gray, 1991 Short dispersed repeats localized in spacer regions of *Chlamydomonas reinhardtii* mitochondrial DNA. Curr. Genet. **19:** 309–312.

Brown, W. M., E. M. Prager, A. Wang and A. C. Wilson, 1982 Mitochondrial DNA sequences of primates: tempo and mode of evolution. J. Mol. Evol. **18:** 225–239.

Bullerwell, C. E., and M. W. Gray, 2004 Evolution of the mitochondrial genome: protist connections to animals, fungi and plants. Curr. Opin. Microbiol. **7:** 528–534.

Cann, R. L., W. M. Brown and A. C. Wilson, 1984 Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. Genetics **106:** 479–499.

Cannone, J. J., S. Subramanian, M. N. Schnare, J. R. Collett, L. M. D'Souza *et al.*, 2002 The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. BMC Bioinformatics **3:** 2.

Cho, Y., J. P. Mower, Y. L. Qiu and J. D. Palmer, 2004 Mitochondrial substitution rates are extraordinarily elevated and variable in a genus of flowering plants. Proc. Natl. Acad. Sci. USA **101:** 17741–17746.

Colleaux, L., M. R. Michel-Wolwertz, R. F. Matagne and B. Dujon, 1990 The apocytochrome b gene of *Chlamydomonas smithii* contains a mobile intron related to both Saccharomyces and Neurospora introns. Mol. Gen. Genet. **223:** 288–296.

Denovan-Wright, E. M., and R. W. Lee, 1994 Comparative structure and genomic organization of the discontinuous mitochondrial ribosomal RNA genes of *Chlamydomonas eugametos* and *Chlamydomonas reinhardtii*. J. Mol. Biol. **241:** 298–311.

Denovan-Wright, E. M., A. M. Nedelcu and R. W. Lee, 1998 Complete sequence of the mitochondrial DNA of *Chlamydomonas eugametos*. Plant Mol. Biol. **36:** 285–295.

Fan, J., and R. W. Lee, 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.

Fan, J., M. N. Schnare and R. W. Lee, 2003 Characterization of fragmented mitochondrial ribosomal RNAs of the colorless green alga *Polytomella parva*. Nucleic Acids Res. **31:** 769–778.

Gaut, B. S., 1998 Molecular clocks and nucleotide substitution rates in higher plants, pp. 93–120 in *Evolutionary Biology*, Vol. 30, edited by M. K. Hecht, R. J. MacIntyre and M. T. Clegg. Plenum Press, New York.

Goldman, N., and Z. Yang, 1994 A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol. Biol. Evol. **11:** 725–736.

Gray, M. W., and P. H. Boer, 1988 Organization and expression of algal (*Chlamydomonas reinhardtii*) mitochondrial DNA. Philos. Trans. R. Soc. Lond. B Biol. Sci. **319:** 135–147.

Gray, M. W., R. Cedergren, Y. Abel and D. Sankoff, 1989 On the evolutionary origin of the plant mitochondrion and its genome. Proc. Natl. Acad. Sci. USA **86:** 2267–2271.

Gray, M. W., B. F. Lang and G. Burger, 2004 Mitochondria of protists. Annu. Rev. Genet. **38:** 477–524.

Hasegawa, M., H. Kishino and T. Yano, 1985 Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. **22:** 160–174.

Keeling, P. J., G. Burger, D. G. Durnford, B. F. Lang, R. W. Lee *et al.*, 2005 The tree of eukaryotes. Trends Ecol. Evol. **20** (12): 670–676.

Kimura, M., 1983 *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge, UK.

Kroymann, J., and K. Zetsche, 1998 The mitochondrial genome of *Chlorogonium elongatum* inferred from the complete sequence. J. Mol. Evol. **47:** 431–440.

Kück, U., K. Jekosch and P. Holzamer, 2000 DNA sequence analysis of the complete mitochondrial genome of the green alga *Scenedesmus obliquus*: evidence for UAG being a leucine and UCA being a non-sense codon. Gene **253:** 13–18.

Kumar, S., K. Tamura and M. Nei, 2004 MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief. Bioinformatics **5:** 150–163.

Laflamme, M., and R. W. Lee, 2003 Mitochondrial genome conformation among CW-group chlorophycean algae. J. Phycol. **39:** 213–220.

Laroche, J., P. Li, L. Maggia and J. Bousquet, 1997 Molecular evolution of angiosperm mitochondrial introns and exons. Proc. Natl. Acad. Sci. USA **94:** 5722–5727.

Lewis, L. A., and R. M. McCourt, 2004 Green algae and the origin of land plants. Am. J. Bot. **91:** 1535–1556.

Li, J. B., S. Lin, H. Jia, H. Wu, B. A. Roe *et al.*, 2003 Analysis of *Chlamydomonas reinhardtii* genome structure using large-scale sequencing of regions on linkage groups I and III. J. Eukaryot. Microbiol. **50:** 145–155.

Lloyd, A. T., and P. M. Sharp, 1992 CODONS: a microcomputer program for codon usage analysis. J. Hered. **83:** 239–240.

Lynch, M., 1996 Mutation accumulation in transfer RNAs: molecular evidence for Muller's ratchet in mitochondrial genomes. Mol. Biol. Evol. **13:** 209–220.

Lynch, M., 2006 Streamlining and simplification of microbial genome architecture. Annu. Rev. Microbiol. **60:** 327–349.

Lynch, M., and J. L. Blanchard, 1998 Deleterious mutation accumulation in organelle genomes. Genetica **102/103:** 29–39.

Lynch, M., B. Koskella and S. Schaack, 2006 Mutation pressure and the evolution of organelle genomic architecture. Science **311:** 1727–1730.

Ma, D. P., Y. T. King, Y. Kim and W. S. Luckett, Jr., 1992 The group I intron of apocytochrome b gene from *Chlamydomonas smithii* encodes a site-specific endonuclease. Plant Mol. Biol. **18:** 1001–1004.

Martin, W., and R. G. Herrmann, 1998 Gene transfer from organelles to the nucleus: How much, what happens, and why? Plant Physiol. **118:** 9–17.

Michaelis, G., C. Vahrenholz and E. Pratje, 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.

Muller, H. J., 1964 The relation of recombination to mutational advance. Mutat. Res. **1:** 2–9.

Muse, S. V., 2000 Examining rates and patterns of nucleotide substitution in plants. Plant Mol. Biol. **42:** 25–43.

Nedelcu, A. M., and R. W. Lee, 1998 A degenerate group II intron in the intronless mitochondrial genome of *Chlamydomonas reinhardtii*: evolutionary implications. Mol. Biol. Evol. **15:** 918–922.

Nedelcu, A. M., R. W. Lee, C. Lemieux, M. W. Gray and G. Burger, 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.

Nei, M., and A. P. Rooney, 2005 Concerted and birth-and-death evolution of multigene families. Annu. Rev. Genet. **39:** 121–152.

Palmer, J. D., K. L. Adams, Y. Cho, C. L. Parkinson, Y. L. Qiu *et al.*, 2000 Dynamic evolution of plant mitochondrial genomes: mobile genes and introns and highly variable mutation rates. Proc. Natl. Acad. Sci. USA **97:** 6960–6966.

Parkinson, C. L., J. P. Mower, Y. L. Qiu, A. J. Shirk, K. Song *et al.*, 2005 Multiple major increases and decreases in mitochondrial substitution rates in the plant family Geraniaceae. BMC Evol. Biol. **5:** 73.

Pesole, G., C. Gissi, A. De Chirico and C. Saccone, 1999 Nucleotide substitution rate of mammalian mitochondrial genomes. J. Mol. Evol. **48:** 427–434.

Pombert, J. F., C. Otis, C. Lemieux and M. Turmel, 2004 The complete mitochondrial DNA sequence of the green alga *Pseudendoclonium akinetum* (Ulvophyceae) highlights distinctive evolutionary trends in the Chlorophyta and suggests a sister-group relationship between the Ulvophyceae and Chlorophyceae. Mol. Biol. Evol. **21** (5): 922–935.

Pombert, J. F., P. Beauchamp, C. Otis, C. Lemieux and M. Turmel, 2006 The complete mitochondrial DNA sequence of the green alga *Oltmannsiellopsis viridis*: evolutionary trends of the mitochondrial genome in the Ulvophyceae. Curr. Genet. **50:** 137–147.

Popescu, C. E., T. Borza, J. P. Bielawski and R. W. Lee, 2006 Evolutionary rates and expression level in Chlamydomonas. Genetics **172:** 1567–1576.

Remacle, C., P. Cardol, N. Coosemans, M. Gaisne and N. Bonnefoy, 2006 High-efficiency biolistic transformation of Chlamydomonas mitochondria can be used to insert mutations in complex I genes. Proc. Natl. Acad. Sci. USA **103:** 4771–4776.

Rocha, E. P., 2004 Codon usage bias from tRNA's point of view: redundancy, specialization, and efficient decoding for translation optimization. Genome Res. **14:** 2279–2286.

Selosse, M., B. Albert and B. Godelle, 2001 Reducing the genome size of organelles favours gene transfer to the nucleus. Trends Ecol. Evol. **16:** 135–141.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin and D. G. Higgins, 1997 The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25:** 4876–4882.

Thorsness, P. E., and T. D. Fox, 1990 Escape of DNA from mitochondria to the nucleus in *Saccharomyces cerevisiae*. Nature **346:** 376–379.

Thorsness, P. E., and T. D. Fox, 1993 Nuclear mutations in *Saccharomyces cerevisiae* that affect the escape of DNA from mitochondria to the nucleus. Genetics **134:** 21–28.

Vahrenholz, C., G. Riemen, E. Pratje, B. Dujon and G. Michaelis, 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.

Wilson, A. C., R. L. Cann, S. M. Carr, M. George, U. B. Gyllensten *et al.*, 1985 Mitochondrial DNA and two perspectives on evolutionary genetics. Biol. J. Linn. Soc. **26:** 375–400.

Wolfe, K. H., W. H. Li and P. M. Sharp, 1987 Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc. Natl. Acad. Sci. USA **84:** 9054–9058.

Wright, F., 1990 The 'effective number of codons' used in a gene. Gene **87:** 23–29.

Xia, X., 2005 Mutation and selection on the anticodon of tRNA genes in vertebrate mitochondrial genomes. Gene **345:** 13–20.

Xia, X., and Z. Xie, 2001 DAMBE: software package for data analysis in molecular biology and evolution. J. Hered. **92:** 371–373.

Yamasaki, T., S. Kurokawa, K. I. Watanabe, K. Ikuta and T. Ohama, 2005 Shared molecular characteristics of successfully transformed mitochondrial genomes in *Chlamydomonas reinhardtii*. Plant Mol. Biol. **58:** 515–527.

Yang, Z., 1997 PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. **13:** 555–556.

Communicating editor: S. Dutcher