Short Communication

Relative rates of evolution among the three genetic compartments of the red alga Porphyra differ from those of green plants and do not correlate with genome architecture

David R. Smith a,*, Jimeng Huab, Robert W. Lee b, Patrick J. Keeling a

a Canadian Institute for Advanced Research, Department of Botany, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4
b Department of Biology, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4R2

ARTICLE INFO

Article history:
Received 8 March 2012
Revised 8 June 2012
Accepted 18 June 2012
Available online 1 July 2012

Keywords:
Genome architecture
Mitochondrial DNA
Mutation rate
Plastid DNA
Substitution rate

ABSTRACT

In photosynthetic eukaryotes, relative silent-site nucleotide substitution rates (which can be used to approximate relative mutation rates) among mitochondrial, plastid, and nuclear genomes (mtDNAs, ptDNAs, and nucDNAs) are estimated to be 1:3:10 respectively for seed plants and roughly equal for green algae. These estimates correlate with certain genome characteristics, such as size and coding density, and have therefore been taken to support a relationship between mutation rate and genome architecture.

Plants and green algae, however, represent a small fraction of the major eukaryotic plastid-bearing lineages. Here, we investigate relative rates of mutation within the model red algal genus Porphyra. In contrast to plants, we find that the levels of silent-site divergence between the Porphyra purpurea and Porphyra umbilicalis mtDNAs are three times that of their ptDNAs and five times that of their nucDNAs. Moreover, relative mutation rates do not correlate with genome architecture: despite an estimated threefold difference in their mutation rate, the mitochondrial and plastid genome coding densities are equivalent—an observation that extends to organisms with secondary red algal plastids. These findings are supported by within-species silent-site polymorphism data from P. purpurea.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

Knowing the mutation rate is important for understanding evolution, but it is a difficult parameter to estimate (Kondrashov and Kondrashov, 2010). Nonetheless, if synonymous and noncoding nucleotide positions (hereafter called silent sites) are assumed to be neutrally evolving, then the silent-site divergence between species can provide an entrée into mutation rate (Kimura, 1983). This approach has some drawbacks: silent sites can be under selective constraints (Andolfatto, 2005; Hershberg and Petrov, 2008), the silent-site divergence can be saturated (Li, 1997), and the number of generations separating two species can be hard to calculate. Most of these problems can, however, be avoided by looking at more than one type of silent site (e.g., synonymous and intergenic sites), comparing only closely related species, and focusing on relative (as opposed to absolute) rates of silent-site substitution, which removes the need for generation- and divergence-time data.

Studies on relative synonymous-site substitution rates (dS) of mitochondrial, plastid, and nuclear DNAs (mtDNAs, ptDNAs, and nucDNAs) from green plants have provided insights into mutation rate and genome evolution (Palmer and Herbon, 1988; Lynch, 2007; Leliaert et al., 2012). For most seed plants, the average dS ratios of mtDNA vs. ptDNA vs. nucDNA are 1:3:10, going up to 1:3:20 in basal angiosperms (Wolfe et al., 1987; Drouin et al., 2008). This contrasts with data from the green algae Chlamydomonas reinhardtii and Chlamydomonas incerta where all three genetic compartments have similar rates of synonymous-site substitution (Popescu and Lee, 2007) and the same is true for Mesostigma viride strains NIES 296 and SAG 50-1 (Hua et al., 2012). These findings, which suggest that seed plants have drastically different mutational patterns than certain green algae, have helped forge the hypothesis that low organelle DNA mutation rates contribute to organelle genome expansion and high rates promote genome contraction (Lynch et al., 2006).

Green algae and land plants represent only a small proportion of the major lineages of plastid-bearing organisms, and outside the green lineage little is known about the relative rates of mutation among mitochondrial, plastid, and nuclear genomes, and how they impact organelle genome complexity. This is because the data needed for these types of analyses were, until recently, difficult to generate, requiring nucleotide sequences from three genetic compartments of two closely related species or strains.

Here, we explore relative rates among the organelle and nuclear genomes of Porphyra—an ancient red algal lineage, comprising...
multicellular marine species, some of which are models for studying the evolution of sex and multicellularity (Mumford and Miura, 1988; Blouin et al., 2011). Porphyra species, and red algae as a whole, have the most gene-rich and architecturally ancestral plastids from all eukaryotes, and both their plastid and mitochondrial genomes are renowned for containing very little noncoding DNA. Moreover, the lateral spread of plastids between distantly related groups means that many diverse eukaryotes harbour red-algal-derived plastids (Teich et al., 2007; Keeling, 2010; Burki et al., 2012). Relative rate data from Porphyra will, therefore, provide a valuable point of comparison with those from other eukaryotic lineages and be useful for addressing hypotheses on organelle genome evolution.

2. Materials and methods

2.1. Specimens

Porphyra umbilicalis (P. um.1, UTEX LB 2951) and P. purpurea were isolated for the US Department of Energy Joint Genome Institute (DOE JGI) Porphyra Genome Project from Schoodic Point, Maine (USA) and Rye, NY (USA), respectively. We also used nucleotide sequence data from an isolate of P. purpurea from Avonport, NS (Canada) (Reith and Munholland, 1995; Burger et al., 1999). Northwestern Atlantic P. umbilicalis reproduces asexually via neutral spores (Blouin et al., 2007) whereas P. purpurea from the Northwest Atlantic normally reproduces via a sexual pathway (Mitman and van der Meer, 1994).

2.2. Organelle and nuclear DNA data

The P. umbilicalis mitochondrial and plastid genomes were assembled using Roche 454 (GS FLX Titanium) and Illumina (Genome Analyzer II) DNA sequence data (GenBank Sequence Read Archive accessions SRX030665-78 and SRX030432-36, respectively), generated by the DOE JGI P. umbilicalis Genome Project and the National Science Foundation Porphyra Genomics Research Collaboration Network (RCN). The Illumina data contained both short (0.3 kilobase; kb) and long (5 kb) insert libraries, allowing us to orient the organelar DNA reads over large genomic distances (e.g., across intergenic regions). The 454 and Illumina reads were mapped onto the P. purpurea organelle genomes (GenBank accessions NC_002007 and NC_000925) (Reith and Munholland, 1995; Burger et al., 1999) with the "Assembly to Reference" program from the Geneious v5.5.6 software suite (Biomatters Ltd., Auckland, New Zealand), using a sensitivity setting of "medium" and a fine tuning setting of "some". Approximately $2 \times 10^7$ P. umbilicalis reads assembled onto the P. purpurea mtDNA and ptDNA, giving complete coverage with >300-fold redundancy. The "mapped" P. umbilicalis reads were reassembled de novo (i.e., without the reference) using the Geneious Assembler (medium sensitivity), giving complete P. umbilicalis mitochondrial and plastid genome sequences (GenBank accessions JQ388471 and JQ408795) (Fig. 1; Supplementary Fig. 1; Supplementary Table S1). To verify that the mtDNA of P. umbilicalis, unlike its P. purpurea counterpart, does indeed lack an inverted repeat, we blasted the P. purpurea mitochondrial inverted repeat element (which also contains the 34 tRNA-coding genes) against all of P. umbilicalis sequencing reads. No hits were found. Also, many of the 454 reads for P. umbilicalis are >0.3 kb, which is longer than the inverted repeat in question, and should therefore have spanned the entire element, preventing assembly errors. P. umbilicalis 454 (GS FLX Titanium) cDNA sequences (GenBank accessions SRX100206-09) confirmed intron locations and the absence of RNA editing in the organelle genomes. Organelle DNA polymorphism data were generated by mapping P. purpurea 454 cDNA sequencing reads (GenBank accessions SRX100229-30) to the P. purpurea mitochondrial and plastid genomes, using the same protocols as above. Nuclear transcripts from P. umbilicalis and P. purpurea were collected from the NoriBLAST databank (Supplementary Table S2) (http://dbdata.rutgers.edu/nori/) – this databank contains filtered, assembled, and annotated EST sequences from P. purpurea and P. umbilicalis, most of which were generated using the 454 cDNA sequences listed above.

2.3. Nucleotide divergence analyses

DNA loci were aligned with MUSCLE (Edgar, 2004), implemented through Geneious, using default settings. Synonymous and nonsynonymous substitutions were measured with the CODEML program of PAML v.4.3 (Yang, 2007), employing the maximum likelihood method (Goldman and Yang, 1994) and the F3 × 4 codon model, and making appropriate adjustments for variation in the genetic code. Substitutions in non-protein-coding regions were estimated with BASEML of PAML, using maximum likelihood and the HKY+G model. Intergenic regions <20 nt were ignored. Nucleotide diversity was calculated with DnaSP v5.10.01 (Librado and Rozas, 2009), using the Jukes and Cantor correction.

3. Results and discussion

3.1. Nucleotide substitution rates between P. purpurea and P. umbilicalis

We measured substitution rates in the organelle and nuclear genomes between Porphyra purpurea and Porphyra umbilicalis (Table 1), two closely related species, distributed primarily throughout the North Atlantic (Brodie et al., 2008; Blouin et al., 2011; Sutherland et al., 2011). Our analyses included complete mitochondrial and plastid genomes and ~65 functionally diverse nuclear loci (Fig. 1; Supplementary Tables S1 and S2). P. purpurea and P. umbilicalis have identical ptDNA gene contents and
datasets (Table 1). The variation in substitution rates were calculated using concatenated and nucDNA (0.43 ± 0.18). An almost identical trend was observed for the mtDNA (1.76 ± 0.58) relative to the ptDNA (0.47 ± 0.22).

The average number of substitutions per synonymous site between the mitochondrial genome than in the plastid and nuclear genomes. The arrangement, including a 5 kb direct repeat, which contains the rRNA-coding genes, but their mitochondrial genomes, although having the same genes, differ in compactness, with the Porphyra purpurea mtDNA containing ~8 kb of noncoding DNA, including five pseudogenes, one intron, and a pair of inverted repeats, that are absent from the P. umbilicalis mtDNA (Supplementary Fig. 1). Recombination between the inverted repeats within the P. purpurea mtDNA is known to promote major genome rearrangements, and because of this the P. purpurea mtDNA is unique among those from red algae for assuming two isomeric conformations (Burger et al., 1999). Our assembly and sequence analyses of the P. umbilicalis mtDNA suggest that it has only one conformation, a finding supported by the lack of inverted repeats in this genome.

Nucleotide substitution rates in Porphyra are higher in the mitochondrial genome than in the plastid and nuclear genomes. The average number of substitutions per synonymous site between the P. purpurea and P. umbilicalis genomes was 1.7–9.4 times greater for the mtDNA (1.76 ± 0.58) relative to the ptDNA (0.47 ± 0.22) and nucDNA (0.43 ± 0.18). An almost identical trend was observed when substitution rates were calculated using concatenated datasets (Table 1). The variation in $d_S$ among the different protein-coding loci within each compartment was 0.79–2.76 (mtDNA), 0.09–2.09 (ptDNA), and 0.03–0.86 (nucDNA) (Supplementary Table S1), which again supports a larger overall silent-site substitution rate in the mitochondrial genome. The rates of substitution at nonsynonymous sites ($d_N/d_S$) were low in all three compartments (<0.1), but were, on average, three and nine times greater for the mtDNA relative to the ptDNA and nucDNA. The $d_N/d_S$ ratio, which can be used to gauge the intensity and directionality of selection, was similar across the three genomes, with averages ranging from 0.04 to 0.06 (Table 1; Supplementary Table S1), which is consistent with purifying selection acting on the nonsynonymous sites. Substitution rates at noncoding sites were also highest in the mtDNA, but for both of the organelle DNAs and the nucDNA they were lower than the respective rates at synonymous sites (Table 1), suggesting that the noncoding regions we employed may be under selective constraints. Overall, these results imply that for P. purpurea and P. umbilicalis the mutation rate of the mitochondrial genome is 3–5 times greater than that of the plastid and nuclear genomes.

3.2.1. Polymorphisms within the P. purpurea organelle genomes

To complement the substitution rate data, we measured the organelle nucleotide diversity between two distinct isolates of P. purpurea, one from Nova Scotia, Canada, and the other from Maine, USA (Table 2). Analyses of complete mitochondrial and almost-complete plastid genome sequences revealed four times more silent-site nucleotide diversity ($\pi_{\text{silent}}$) in the mtDNA than the ptDNA (3.6 × 10⁻³ vs. 0.7 × 10⁻³). A previous report on the P. purpurea mitochondrial genome noted similar levels of silent-site mtDNA diversity (~3.5 × 10⁻³) (Burger et al., 1999). Nucleotide diversity at nonsynonymous sites was also four times greater in the mitochondrial compartment relative to that of the plastid (1.0 × 10⁻³ vs. 0.24 × 10⁻³).

Although influenced by many forces, $\pi_{\text{silent}}$ can provide a proxy for 2N_{eff} twice the effective number of gene copies at a locus times the per-generation, per-site mutation rate (Nei, 1987; Lynch, 2007). Given that both the mtDNA and ptDNA of Porphyra are unparentally inherited – with rare occurrences of biparental inheritance for both mtDNA and ptDNA (Choi et al., 2008; Niwa et al., 2010) – one may expect them to have a similar N_{eff}. If true, the difference in $\pi_{\text{silent}}$ between the P. purpurea organelle genomes could be explained by a 4-fold higher mutation rate in the mtDNA relative to the ptDNA, thus, supporting the findings of the substitution rate analyses described above.

3.3.1. Relative mutation rates in Porphyra: the opposite of land plants

The substitution rate and nucleotide diversity data indicate that for P. purpurea and P. umbilicalis the mutation rate of the mitochondrial genome exceeds that of its plastid and nuclear counterparts. To the best of our knowledge, this is the first investigation on relative mutation rates among mtDNA, ptDNA, and nucDNA within the red algae. There has been a large and impressive effort to measure nucleotide sequence variation within and between red algal lineages (Saunders and Hommersand, 2004), including Porphyra (Lindstrom and Fredericq, 2003; Milstein et al., 2008; Teasdale and Klein, 2010). These studies have excelled at resolving phylogenetic relationships among red algae (Verbruggen et al., 2010; Sutherland et al., 2011), but because most use a single locus from only one or two genetic compartments they could not address the relative rates of mutation between compartments. Nevertheless, most of the studies employing both mitochondrial and plastid loci observed more inter- and intra-specific nucleotide divergence in the mtDNA than the ptDNA, particularly for lineages in the Bangiales, like Porphyra (Robba et al., 2006; Milstein et al., 2008), which is in line with the results presented here. Some, however, have found approximately equal or slightly greater divergence in nucDNA than mtDNA or ptDNA. Teasdale and Klein (2010) showed that for various P. umbilicalis isolates there was no variation in a 0.3 kb stretch of noncoding mtDNA, but that there was 0–3.3% variation for a 0.7 kb noncoding nucDNA region. Klein et al. (2003) found the interspecific divergence among various species of Porphyra,
Table 2
Polymorphisms within the *Porphyra purpurea* mitochondrial and plastid genomes.

<table>
<thead>
<tr>
<th></th>
<th>mtDNA</th>
<th>ptDNA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>S</td>
</tr>
<tr>
<td><strong>Ribosomal DNA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synonymous sites*</td>
<td>3995</td>
<td>5</td>
</tr>
<tr>
<td>Nonsynonymous sites</td>
<td>4944</td>
<td>19</td>
</tr>
<tr>
<td>Introns</td>
<td>16,521</td>
<td>17</td>
</tr>
<tr>
<td>Intergenic sites</td>
<td>1581</td>
<td>3</td>
</tr>
<tr>
<td>Silent sites†</td>
<td>7183</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>13,708</td>
<td>49</td>
</tr>
</tbody>
</table>

N, number of sites (includes all sites in the nucleotide alignment, including those with indels); S, number of polymorphic sites; Indels, insertion–deletion events (indels involving more than one nucleotide are counted as a single event); π, nucleotide diversity; n/a, not applicable.

* Includes large- and small-subunit ribosomal RNA-coding regions.
‡ Includes synonymous, intergenic and (when applicable) intronic sites.

including *P. purpurea* and *P. umbilicata*, was 4–14% for the nuclear SSU rRNA gene and 3–11% for the ptDNA-encoded *rbcL* gene.

If, indeed, the relative mutation rates of *Porphyra* are greater in the mtDNA as compared to the ptDNA and nucDNA, it would represent a marked difference from what is observed in green plants (Table 3). For both *Chlamydomonas* and *Mesostigma*, which together represent the two main green algal lineages (the chlorophytes and charophytes, respectively), the rates of silent-site substitution are almost constant (within a factor of two) across the three genetic compartments (Table 3). For seed plants, with few exceptions (Mower et al., 2007; Sloan et al., 2008), the silent-site substitution rates of mtDNA are, on average, 3- and 10-fold lower than those of the ptDNA and nucDNA (Table 3) – the opposite trend to that of *Porphyra*. There is reason to believe, however, that the relative rates of *Porphyra* organelle genomes may be similar to those of species with secondary plastids derived from red algae. In a comparison of a haptophyte, a stramenopile, and an apicomplexan, the rates of silent-site mtDNA divergence greatly exceeded those of the ptDNA (Smith and Keeling, 2012), paralleling the trend for *Porphyra*, which has a primary red plastid. Altogether, these data suggest that the relative rates of mutation for mtDNA, ptDNA, and nucDNA differ widely across the eukaryotic domain, and that a mtDNA(ptDNA) mutation rate ratio of ≥1 may be the norm among plastid-bearing protists, with a ratio of <1 being restricted to seed plants.

### 3.4. Mutation rate and organelle genome architecture

Certain studies have found a positive relationship between mutation rate and coding density within organelle genomes (Lynch et al., 2006; Smith and Lee, 2010) (but see Sloan et al., 2012), and references therein, which note the opposite trend. One explanation for these observations is the mutational burden (or hazard) hypothesis (MBH) (Lynch, 2007), which argues that a higher mutation rate makes for a less permissive environment for the proliferation of noncoding nucleotides, such as repeats and introns. Data on the relative mutation rates of organelle genomes allow for a reasonable test of this hypothesis, at least in these systems. The MBH would predict that the relative silent-site substitution rates of mtDNA and ptDNA should reflect the relative proportion of coding and noncoding nucleotides within these genomes, where a lower relative rate is associated with a larger relative noncoding DNA density (i.e., smaller dS = more bloated genome) and vice versa. When looking at average values within major groups (Table 3; Supplementary Table S3), the data from green plants support the MBH but those from *Porphyra* and species with secondary red plastids do not (Fig. 2). For seed plants, the mtDNA has on average a smaller dS and a larger noncoding DNA density than the ptDNA, and in green algae the mtDNA and ptDNA have similar average dS values and noncoding densities (Fig. 2). Conversely, for *Porphyra* and certain species with secondary red algal-derived plastids, a much larger dS in the mtDNA is associated with either a slight increase or no change in noncoding density relative to the ptDNA (Fig. 2), which is inconsistent with the MBH. That said, more than half of the noncoding nucleotides within the *Porphyra purpurea* mtDNA appear to have been acquired through a recent lateral DNA transfer event, potentially involving a mitochondrial plasmid (Burger et al., 1999). Moreover, the *P. umbilicata* mtDNA, which does not harbour plasmid-like sequences (Fig. 1; Supplementary Fig. S1), is more compact than the ptDNA (89% vs. 86% coding).

Table 3
Relative synonymous substitution rates among mitochondrial, plastid, and nuclear genomes within various eukaryotic lineages.

<table>
<thead>
<tr>
<th>Substitutions per synonymous site</th>
<th>Substitution rate ratios (mt:pt:nuc)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mtDNA</td>
</tr>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td><strong>Red algae</strong></td>
<td></td>
</tr>
<tr>
<td><em>Porphyra</em></td>
<td>1.76 (0.58)</td>
</tr>
<tr>
<td><strong>Green algae</strong></td>
<td></td>
</tr>
<tr>
<td><em>Chlamydomonas</em></td>
<td>0.29 (0.05)</td>
</tr>
<tr>
<td><em>Mesostigma</em></td>
<td>0.17 (0.11)</td>
</tr>
<tr>
<td><strong>Seed plants</strong></td>
<td></td>
</tr>
<tr>
<td>Angiosperms</td>
<td>0.13 (0.01)</td>
</tr>
<tr>
<td>Gymnosperms</td>
<td>0.28 (0.02)</td>
</tr>
</tbody>
</table>

Synonymous-site substitution rates are based on averages among loci, not concatenations.

* Data from Popescu and Lee (2007) and Hua et al. (2012).
† Data from Drouin et al. (2008).
When ignoring noncoding density, some features of mtDNA architecture within *Porphyra* and taxa with secondary red plastids are compatible with genome evolution under the MBH: an alternative genetic code, reduced gene content, and smaller genome size in the mtDNA relative to the ptDNA are all consistent with a higher relative rate in the mitochondrial compartment (Lynch et al., 2006; Lynch, 2007).

As more data on relative rates of evolution become available, our understanding of the mutational patterns among mtDNAs, ptDNAs, and nucDNAs across major groups will improve, which will allow us to test how much variation in eukaryotic genome evolution can be explained by the MBH, and how much is derived through other forces. In regards to the data presented here, it will be particularly interesting to see how the relative substitution rates across the genetic compartments of *Porphyra* (and other red algae) compare to those of cryptophyte algae, which have four different genomes, including a red-algal-derived plastid and a highly reduced and compacted nucleomorph, derived from a red algal nucleus.

**Acknowledgments**

We are grateful to the National Science Foundation (NSF) for their support of the *Porphyra* Research Collaboration Network (David Smith, RCN Member) and for access to the *P. umbilicalis* and *P. purpurea* genome and transcriptome sequencing data, which were generated with the help of the following grants (and PIs): NSF 0741907 (Susan Brawley, Elisabeth Gantt, Arthur Grossman, John Stiller), National Oceanic and Atmospheric Administration (NOAA) NAO600AR4170108 (S. Brawley), NOAA R/A-38 (Charles Yarish), NSF EF-0626678 (Senjie Lin), DOE JGI Contract DE-AC02-05CH11231 (S. Brawley, E. Gantt, A. Grossman, J. Stiller), NSF 0732440 and Division of Environmental Biology 1004213 (Debashish Bhattacharya). Jimeng Hua is supported by a grant to Robert Lee from the Natural Sciences and Engineering Research Council of Canada (NSERC). Patrick Keeling is funded by a grant from NSERC (227301) and is a Fellow of the Canadian Institute for Advanced Research and a Senior Scholar of the Michael Smith Foundation for Health Research. D. Smith is supported by postdoctoral fellowships from NSERC and the Izaak Walton Killam Trusts.

**Appendix A. Supplementary material**

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2012.06.017.

**References**


