Complete mitogenome sequence of *Ricasolia amplissima* (Lobariaceae) reveals extensive mitochondrial DNA rearrangement within the Peltigerales (lichenized ascomycetes)

Author(s): Antoine Simon, Yang Liu, Emmanuël Sérusiaux and Bernard Goffinet
Source: The Bryologist, 120(3):335-339.
Published By: The American Bryological and Lichenological Society, Inc.
https://doi.org/10.1639/0007-2745-120.3.335
URL: http://www.bioone.org/doi/full/10.1639/0007-2745-120.3.335

Complete mitogenome sequence of *Ricasolia amplissima* (Lobariaceae) reveals extensive mitochondrial DNA rearrangement within the Peltigerales (lichenized ascomycetes)

Antoine Simon[1], Yang Liu[2,3], Emmanuël Sérusiaux[1] and Bernard Goffinet[2,4]

[1] *Evolution and Conservation Biology, InBIOS research center, University of Liège, Quartier Vallée 1, Chemin de la Vallée 4, B-4000 Liège, Belgium;* [2] *Department of Ecology and Evolutionary Biology, 75 North Eagleville Road, Storrs, CT 06269, U.S.A.;* [3] *Fairy Lake Botanical Garden, Shenzhen & Chinese Academy of Sciences, China*

Abstract. The structure of mitochondrial genomes varies among non-lichenized fungi in terms of their genic and intronic content and genic order. Whether lichenized fungal mitogenomes are equally labile is unknown due to the paucity of available mitogenomes. We assembled the mitogenome of *Ricasolia amplissima* (Peltigerales, Lobariaceae), using massive parallel sequencing, and compared its structure to that of two species of *Peltigera* (Peltigeraceae). The mitochondrial genome of *R. amplissima* comprised 82,333 bp, with a 29.8% G+C content, and holds 15 unique protein-coding genes, 29 tRNA genes, two rRNA genes, and one non-coding RNA gene. Although the protein-coding gene content in the mitogenome of *Peltigera* and *Ricasolia* was identical, the relative gene order differed substantially, revealing that significant gene rearrangements also characterize the evolution of mitogenomes of lichenized ascomycetes at a relatively shallow phylogenetic depth, such as within the order Peltigerales.

Keywords. Gene duplication, genome rearrangement, Lecanoromycetes, lichen, Peltigerinae.

◆ ◆ ◆

Mitochondrial (mt) genomes consist of a single chromosome in fungi (Aguileta et al. 2014), with the exception of a species of Glomeromycota (Nadimi et al. 2015). Fungal mitogenomes vary extensively among and within major fungal orders, both in terms of size [e.g., 20,063 bp for *Candida glabrata* (Saccharomycetes) to 235,849 bp for *Rhizoctonia solani* (Basidiomycetes)] and gene content/order (Aguileta et al. 2014; Koszul et al. 2003; Losada et al. 2014). Whether this high variability in composition and structure also occurs among closely related families of lichenized fungi is unknown. Indeed, in spite of the fact that one-in-five fungal species is lichenized (Honegger 2008), the mitogenomes of only two lichen-forming ascomycetes have been published to date: *Peltigera malacea* (JN088164) and *P. membranacea* (JN088165; Xavier et al. 2012). Their mtDNA are collinear and mainly differ in the number of group I introns and the presence of a non-functional *dpoB* gene in *P. membranacea* (apparently functional in *P. malacea*).

Peltigeralean lichen-forming fungi form a well-defined order within the Lecanoromycetes, the largest class of almost exclusively lichenized fungi and third most diverse in the fungal kingdom (Kirk et al. 2008; Lücking et al. 2016; Miadlikowska et al. 2014). The Peltigerales arose near the Jurassic/Cretaceous boundary, i.e. about 142–150 mya (Prieto & Wedin 2013). Its species are accommodated among ten families (Spribille & Muggia 2013), including the Lobariaceae and the Peltigeraceae, which are the focus of this paper.

The aim of the present genome report is to assess the architectural stability of the fungal mt chromosome among species from two closely related families. We assembled the mitogenome of *Ricasolia amplissima* (Scop.) De Not. (Lobariaceae, Peltigerales) using a massive parallel sequencing approach, and compared its gene architecture to that of *Peltigera* (Peltigeraceae, Peltigerales; Xavier et al.

[4] Corresponding author's e-mail: bernard.goffinet@uconn.edu

DOI: 10.1639/0007-2745-120.3.335

0007-2745/17/$0.65/0

2012), i.e. the most closely related genus whose mitochondrial genome has been published.

Material & Methods

Total DNA was extracted from a specimen of *Ricasolia amplissima* free of any visible symptoms of parasitism [NORWAY. Hordaland: 2015, *Tønsberg 44719* (conn)]. DNA was isolated using the NucleoSpin® Plant II Midi kit (Macherey-Nagel, Bethlehem, PA, USA) following the manufacturer's guidelines for three different thallus fragments: two wherein the fungus is primarily associated with a cyanobacterium (i.e. the so-called cepholodia) and one wherein it is primarily associated with a green alga (NCBI short read archive BioSample accessions SAMN06251167, SAMN06251171, and SAMN06251162, respectively; e.g. see Tønsberg & Goward 2001 for description of this tripartite association). Three Nextera (Illumina, San Diego, CA, U.S.A.) genomic libraries were constructed (i.e., one for each DNA extraction). Two libraries were multiplexed and sequenced with two unrelated libraries; a third library was multiplexed and sequenced with three other unrelated libraries. Each sequencing run used an Illumina MiSeq instrument with a 600-cycle v3 sequencing kit (Illumina, San Diego, CA, U.S.A.).

The reads were filtered and trimmed via Trimmomatic v0.32 (Bolger et al. 2014). Following this trimming step, totals of 9,636,857 (SRA accession: SAMN06251162), 14,036,794 (SRA accession: SAMN06251167), and 10,804,290 (SRA accession: SAMN06251171) reads were obtained. Filtered, trimmed paired-end reads were *de novo* assembled independently in CLC genomics workbench 6.5 (CLC bio, Aarhus, Denmark) using the following parameters: mismatch cost = 2, insertion cost = 3, deletion cost = 3, length fraction = 0.8, similarity fraction = 0.95. Each assembly resulted in 11 similar contigs identified by BLAST as fragments of the *R. amplissima* mt genome, and thus subsequent mappings were performed onto the contigs from one single assembly (SRA accession: SAMN06251167), which had the largest assembly size. In one instance, one of these contigs was significantly longer in another assembly (SRA accession: SAMN06251171) and was consequently favored. These eleven contigs were imported in Geneious 7.1.3 (Biomatters Ltd., Auckland, New Zealand); the reads from all libraries were mapped iteratively on the recovered contigs (following the methodology used by Fucíková et al. 2014) to extend and potentially connect them. Gaps between the resulting four contigs were filled via polymerase chain reaction (PCR) and Sanger sequencing (see annotation in GenBank accession). The sequence ambiguity of the duplicated *nad*1 gene was resolved with the same approach (see annotation in GenBank accession). The resulting sequence was validated by mapping the reads against it, confirming that sequencing depth was low only for the zones recovered by Sanger sequencing (**Supplementary Fig. S1**). When the reads SAMN06251167 were mapped against the resulting 82,333 bp-long mt genome (using the medium-low sensitivity settings in Geneious), the depth of reads averaged 269.2 ×.

The mtDNA of *R. amplissima* was annotated in Geneious by mapping the mt genes of *P. membranacea* and *P. malacea* (JN088165, JN088164) to it, and by using tRNAscan SE 1.21 (Lowe & Eddy 1997). Annotation accuracy for protein coding genes was examined by BLASTX comparisons (Altschul et al. 1990). The *R. amplissima* mitogenome was deposited in GenBank under the accession number KY853549. Raw DNA sequence reads were deposited at the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under BioProject PRJNA362948.

Results and Discussion

The circular mt genome of *R. amplissima* comprised 82,333 bp, with a 29.8% G+C content. The genome contained six introns for a cumulative size of 7,479 bp, whereas exons, non-coding RNA genes, and spacers amounted to 15,926 bp, 7,188 bp, and 51,740 bp, respectively. The mtDNA encoded 49 genes: 15 unique protein-coding genes, 29 tRNA genes (which covered all 20 common amino acids), two rRNA genes, and one non-coding RNA gene involved in the Ribonuclease P (**Fig. 1**). Duplicated gene copies were identified for *nad1* and *atp9*. Indeed, the mitogenome of *R. amplissima* included two fragments with homology to *nad1*. One copy (annotated *nad1*) occurred upstream of the small subunit rRNA and likely encoded a 312-aa long protein, whereas the other (annotated *nad1** on **Fig. 1**) occurred upstream of *cox1* and lacked a canonical stop codon. The two copies were 99% identical over a 936-bp long coding sequence. The exact sequence identity of each copy could be recovered by Sanger sequencing one copy to determine the linkage of bases in polymorphic sites. The gene *atp9* also

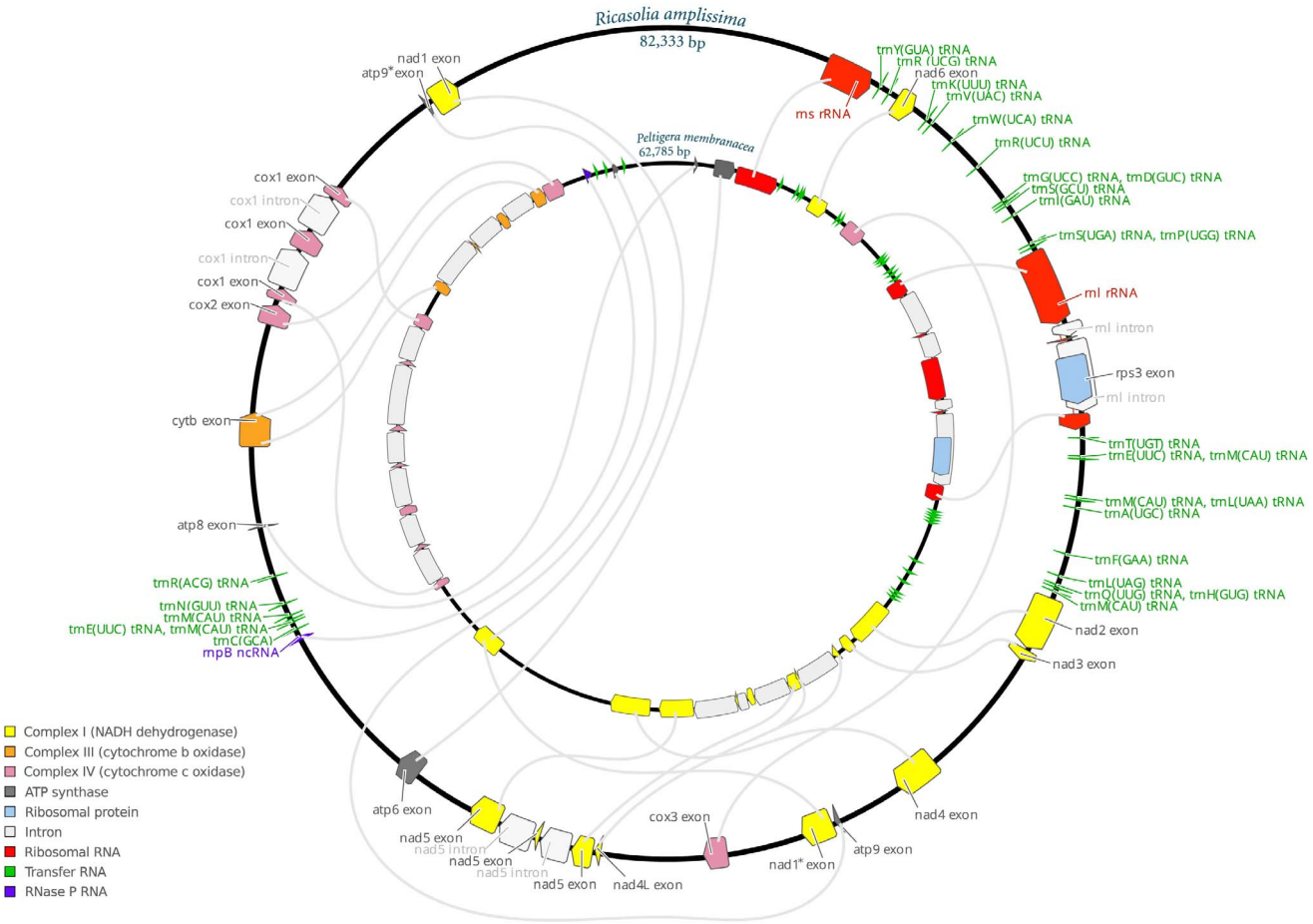

Figure 1. Comparison of the architecture (genic and intronic distribution) of the mitogenome of *Ricasolia amplissima* (outer circle) and *Peltigera membranacea* (inner circle; JN088165; Xavier et al. 2012). All genes are transcribed clockwise. Correspondences between genomes are shown by gray curves. The two genomes are not represented at the same scale.

appeared to have been duplicated; one functional copy (annotated *atp9*) occurred upstream of *nad1**, comprised 225 bp and likely encoded a 75-aa long protein, whereas the other (annotated *atp9** on **Fig. 1**) occurred upstream of *nad1* and was treated as a pseudogene owing to its disrupted reading frame. The two copies revealed 47% pairwise identity over 225 bp. Although rare in fungal mitogenomes, gene duplications have been described, including for the *atp9* gene (Mardanov et al. 2014).

Gene delimitation was unambiguous in *R. amplissima*, with the following exceptions: no canonical start codon was identified for *atp6*, *cox2* and *cytb*, and no canonical stop codon was detected for *cox1*, *cox2*, *cytb*, *nad1**, *nad3* and *nad4*. Xavier et al. (2012) also reported the lack of a stop codon for *cytb* and a start for *cox2*. In the case of *nad6*, we considered the alternative start codon GTG as the beginning of the gene. GTG is also used as start codon in mt genes of some early land plant lineages, such as liverworts and mosses (Liu et al. 2011). In other genes, when the delimitation was ambiguous due to the absence of start/stop codons, the limits were demarked so that they corresponded to the length of annotated genes in *Peltigera*; additional exons for these genes were not detected in the surrounding intergenic spacer regions using the BLASTX tool.

The mitogenome of *R. amplissima* (82,333 bp) was much larger than that of *P. malacea* (63,363 bp) and *P. membranacea* (62,785 bp). It also differed in the number and length of introns (six in *R. amplissima*, 19 in *P. malacea*, 17 in *P. membranacea*). As in the mitogenome of *P. malacea* and *P. membranacea*, tRNA genes were clustered in *R. amplissima*. The two major clusters, comprising 12 and 11 tRNA genes, were located upstream and downstream of the *rnl* gene, respectively, as in the *Peltigera* mt genomes. The third cluster occurred between *rnpB* and *atp8* in both *R.*

Table 1. Characterization of the mitogenome of *Ricasolia amplissima*, *Peltigera malacea* and *P. membranacea*.

	Ricasolia amplissima KY853549	*Peltigera malacea* JN088164	*Peltigera membranacea* JN088165
Assembly size (bp)	82,333	63,363	62,785
Exons, cumulative size (bp)	15,926	14,805	14,862
Introns, cumulative size (bp)	7,479	24,953	22,730
Non-coding RNA genes, cumulative size (bp)	7,188	6,792	6,796
Spacers, cumulative size (bp)	51,740	16,813	18,397
G+C (%)	29.8	27.2	27.1
Unique protein-coding genes	15	16	15
Average number of exons per gene	1.267	1.875	1.867
tRNAs	29	26	26
cox1 introns	2	9	6

amplissima and in *Peltigera* spp. but included six versus only three tRNAs. Altogether, the mtDNA of *R. amplissima* comprised 29 tRNA genes (vs. 26 in *Peltigera*). More genomic information for these three mitogenomes is shown in **Table 1**.

Whereas protein-coding gene content in *R. amplissima* and *Peltigera* spp. was identical, the relative gene order differed substantially (**Fig. 1**), indicating massive gene order rearrangements in the mitogenomes of lichenized ascomycetes, even at a relatively shallow phylogenetic depth, such as within the order Peltigerales. In fungal and plant lineages variation in mt gene order is linked to the presence of repeated sequences scattered among intergenic regions, which enable intragenomic recombinations (e.g., Aguileta et al. 2014; Liu et al. 2014). The mitogenome of *R. amplissima* comprised 25 pairs and four triplets of direct or inverted repeats greater than 150 bp sharing a sequence identity higher than 85%. It also included six tandem repeats greater than 50 bp (**Supplementary Fig. S2, Supplementary Table S3**; see also Xavier et al. 2012 for *Peltigera* spp.). In the Ascomycota, the extreme gene order variability observed here contrasts the high structural genome stability detected, for example, within the Hypocreales and the Onygenales, but follows a pattern similar to that previously observed for the Saccharomycetales (Aguileta et al. 2014). The longest stretch of collinearity between the mtDNA of *R. amplissima* and *Peltigera* spp. covered only three protein-coding genes (*rps3*, *nad2* and *nad3*), indicating that genomic structure in Peltigerales is extremely labile in a phylogenetic context.

Acknowledgments

This project is partially funded by the National Science Foundation through grant DEB–1354631 to BG. We thank Tor Tønsberg for collecting fresh material of the studied species. We would like to acknowledge Laurent Gohy and Rafael Medina for assistance with laboratory work at the University of Liège and the University of Connecticut, respectively. We thank two reviewers who provided valuable comments that helped improve the manuscript. AS is a PhD student at the University of Liège and acknowledges the financial support by FRIA, a grant of the Belgian Research Foundation (F.R.S.-FNRS).

Literature Cited

Aguileta, G., D. M. De Vienne, O. N. Ross, M. E. Hood, T. Giraud, E. Petit & T. Gabaldón. 2014. High variability of mitochondrial gene order among fungi. Genome Biology and Evolution 6: 451–465.

Altschul, S., W. Gish & W. Miller. 1990. Basic Local Alignment Search Tool. Journal of Molecular Biology 215: 403–410.

Bolger, A. M., M. Lohse & B. Usadel. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30: 2114–2120.

Fucíková, K., P. O. Lewis, D. González-Halphen & L. A. Lewis. 2014. Gene arrangement convergence, diverse intron content, and genetic code modifications in mitochondrial genomes of Sphaeropleales (Chlorophyta). Genome Biology and Evolution 6: 2170–2180.

Honegger, R. 2008. Mycobionts. Pages 27–39. In: T. H. Nash III(ed.),Lichen Biology. 2nd Edition. Cambridge University Press, Cambridge.

Koszul, R., A. Malpertuy, , L. Frangeul, C. Bouchier, P. Wincker, A. Thierry, S. Duthoy, et al. 2003. The complete mitochondrial genome sequence of the pathogenic yeast *Candida* (*Torulopsis*) *glabrata*. FEBS Letters 534: 39–48.

Kirk, P. M., P. F. Cannon, D. W. Minter & J. A. Stalpers. 2008. Dictionary of the Fungi. 10th Edition. CAB International, Wallingford.

Liu, Y., J. Y. Xue, B. Wang, L. Li & Y. L. Qiu. 2011. The mitochondrial genomes of the early land plants *Treubia lacunosa* and *Anomodon rugelii*: Dynamic and conservative evolution. PLoS ONE 6: e25836.

Liu, Y., R. Medina & B. Goffinet. 2014. 350 my of mitochondrial genome stasis in mosses, an early land plant lineage. Molecular Biology and Evolution 31: 2586–2591.

Losada, L., S. B. Pakala, N. D. Fedorova, V. Joardar, S. A. Shabalina, J. Hostetler, S. M. Pakala, et al. 2014. Mobile elements and mitochondrial genome expansion in the soil fungus and potato pathogen *Rhizoctonia solani* AG-3. FEMS Microbiology Letters 352: 165–173.

Lowe, T. M. & S. R. Eddy. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Research 25: 955–964.

Lücking, R., B. P. Hodkinson & S. D. Leavitt. 2016 [2017]. The 2016 classification of lichenized fungi in the Ascomycota and Basidiomycota–Approaching one thousand genera. The Bryologist 119: 361–416.

Mardanov, A. V., A. V. Beletsky, V. V. Kadnikov, A. N. Ignatov & N. V. Ravin. 2014. The 203 kbp mitochondrial genome of the phytopathogenic fungus *Sclerotinia borealis* reveals multiple invasions of introns and genomic duplications. PLoS ONE 9: e107536.

Miadlikowska, J., F. Kauff, F. Högnabba, J. C. Oliver, K. Molnár, E. Fraker, E. Gaya, et al. 2014. A multigene phylogenetic synthesis for the class Lecanoromycetes (Ascomycota): 1307 fungi representing 1139 infrageneric taxa, 317 genera and 66 families. Molecular Phylogenetics and Evolution 79:132–168.

Prieto, M. & M. Wedin. 2013. Dating the diversification of the major lineages of Ascomycota (Fungi). PLoS ONE 8: e65576.

Spribille, T. & L. Muggia. 2013. Expanded taxon sampling disentangles evolutionary relationships and reveals a new family in Peltigerales (Lecanoromycetidae, Ascomycota). Fungal Diversity 58: 171–184.

Tønsberg, T. & T. Goward. 2001. *Sticta oroborealis* sp. nov. and other Pacific North American lichens forming dendriscocauloid cyanotypes. The Bryologist 104: 12–23.

Xavier, B. B., , V. P. W. Miao, Z. O. Jónsson & Ó. S. Andrésson. 2012. Mitochondrial genomes from the lichenized fungi *Peltigera membranacea* and *Peltigera malacea*: features and phylogeny. Fungal Biology 116: 802–814.

manuscript received February 14, 2017; accepted July 13, 2017.

Supplementary documents online:

Supplementary Fig. S1. Read depth of *Ricasolia amplissima* obtained after mapping the reads SAMN06251167 (using the medium-low sensitivity settings in Geneious).

Supplementary Fig. S2. Distribution of repeats in the mitogenome of *Ricasolia amplissima*. Tandem repeats greater than 50 bp are shown in gray inside the circular chromosome. Repeats greater than 150 bp with sequence identity higher than 85% are shown inward, with sets of identical repeats color-coded.

Supplementary Table S3. Location of repeats in the mitogenome of *Ricasolia amplissima* (KY853549). Tandem repeats included are greater than 50 bp; direct or inverted repeats considered are greater than 150 bp and share sequence identity higher than 85%.