

A novel lineage of non-photosynthetic chlamydomonadalean algae with a peculiar plastid genome.

K. Záhonová¹, T. Pánek¹, S. C. Treitli², N. Yubuki³, E. Zadrobílková⁴, V. Yurchenko¹, I. Čepička⁴, M. Eliáš¹

1. Life Science Research Centre, Department of Biology nad Ecology, Faculty of Science, University of Ostrava, Ostrava, Czech Republic

2. Department of Parasitology, Faculty of Science, Charles University in Prague, Prague, Czech Republic

3. Departments of Zoology and Botany, University of British Columbia, Vancouver, Canada

4. Department of Zoology, Faculty of Science, Charles University in Prague, Prague, Czech Republic

The biflagellate *Polytoma* and the quadriflagellate *Polytomella* are unrelated colourless genera of green algal order Chlamydomonadales. Nominal *Polytoma* species are diphyetic, indicating at least three independent losses of photosynthesis among Chlamydomonadales. We isolated and investigated two new *Polytoma*-like strains (AMAZONIE and MBURUCU). Based on 18S rDNA gene sequences they turned out to represent a separate deep clade and the fourth independently arisen non-photosynthetic chlamydomonadalean lineage. Morphological differences together with several compensatory base changes in the ITS2 rDNA region indicate that AMAZONIE and MBURUCU represent two different species. Using the Illumina sequencing we obtained genomic and transcriptomic data from the AMAZONIE strain. Searching for plastid genome sequences returned 34 genes encoding plastid proteins, none of them with a photosynthesis-specific function. The genes were distributed on separate contigs with each gene flanked by repetitive non-coding regions. The plastid genome of the AMAZONIE strain may thus be organized similarly to the recently characterized inflated repeat-rich plastid genome of *Polytoma uvella*, or it may alternatively consist of individual single-gene “minichromosomes” with the terminal repeats serving as telomeres. To distinguish between these possibilities we generated long-read sequencing data using the Oxford Nanopore technology. The yield of plastid genome-derived data was low due to heavy bacterial contamination of the culture, but enabled us to assemble a contig of 22,875 bp containing two protein-coding genes (*rpoC1* and *atpB*) and two tRNA genes, providing the first evidence that the AMAZONIE plastid genome may potentially be unsegmented. The length of non-coding repeat-rich regions separating the genes in the contig suggests that the whole genome may be even more inflated than that of *P. uvella*. Further sequencing (including an Illumina mate pair library) and an analysis of the transcriptome are underway to improve the plastid genome assembly and to illuminate the function of the AMAZONIE plastid.