

Mitochondrial genome of strain SRT308: insights into structure and gene content of the ancestral euglenozoan mitochondrial genome.

E. Yazaki¹, T. Shiratori¹, K. Kume², T. Hashimoto^{1,3}, K. Ishida¹ and Y. Inagaki³

1. Faculty of Life and Environmental Sciences, University of Tsukuba, Ibaraki, Japan

2. Graduate School of Life and Environmental Sciences, University of Tsukuba, Ibaraki, Japan

3. Center for Computational Science, University of Tsukuba, Ibaraki, Japan

Euglenozoa is a large protist assemblage comprising three major subgroups, namely Euglenida, Diplonemea and Kinetoplastea. Members of this assemblage are known for a large diversity in mitochondrial (mt) genome structure. The mt genome of *Euglena gracilis*, a representative of euglenids, are likely composed of multiple linear chromosomes, which harbor only 6 protein-coding genes and fragmented ribosomal RNA genes. No RNA editing has been detected in the transcripts from *Euglena* mt genome. Diplonemid mitochondrial genomes were found to comprise multiple minicircular chromosomes, each of which carries a partial gene fragment, and initial transcripts from the mt minicircles receive both *trans*-splicing and RNA editing to yield mature mRNAs. Mitochondria of members of Kinetoplastea contain a complex network of two types of circular chromosomes, maxicircles and minicircles, and mt gene transcripts are edited intensively prior to translation. To understand how the complex genome structures and RNA editing mechanisms observed among the extant euglenozoans emerged, it is important to infer the mt genome structure of the ancestral euglenozoan. The aforementioned issue can be addressed by analyzing the mt genome of a novel protist strain SRT308, which was placed robustly at the base of the Euglenozoan clade in our phylogenomic analysis based on 153 nucleus-encoded gene sequences. We successfully reconstructed the complete SRT308 mt genome of 61 Kb in length in this study. In this mt genome, we identified 35 open reading frames (ORFs) and 29 structural RNA genes (2 and 27 are rRNA and tRNA genes, respectively). 19 out of the 35 ORFs are functionally assignable, and some of them were not found in *Euglena*, *Diplonema* or kinetoplastids. The SRT308 mt genome suggests that this protist (and the ancestral euglenozoan as well) possesses the mt genome larger in both size and gene content than the euglenozoan mt genomes studied to date.