

## Comparative analysis of malawimonad and jakobid nuclear genomes

**M. Sarrasin, R. Derelle, G. Burger and BF. Lang**

*Department of Biochemistry, Robert-Cedergren Centre of Bioinformatics and Genomics, Université de Montréal, Montréal, QC, H3C 3J7, Canada*

Jakobids and malawimonads are two distinct lineages of bacterivorous unicellular flagellates that diverge close to the eukaryotic origin. While jakobids are clearly part of the Discoba ('excavates'), the position of malawimonads remains controversial, rather associating with the Amoebozoa/Opisthokonta assembly of eukaryotes. In order to obtain more detailed information on the coding capacity and gene structure of their nuclear genomes, we have developed bioinformatics procedures for genome assembly that handle substantial contamination of eukaryotic reads with that from food bacteria. We succeeded with high quality assemblies of the jakobids *Reclinomonas americana*, *Jakoba bahamensis*, *Jakoba libera*, *Seculamonas ecuadoriensis* and *Andalucia godoyi* (the latter in collaboration with M. Elias), and the diverse malawimonads *Malawimonas californiana*, *M. jakobiformis* and *M. sp.* We have further sequenced transcriptomes from these species to facilitate inference of gene models. As there are no genome sequences of closely related protists to assist genome annotation, we have modified popular bioinformatics pipelines to improve the accuracy of gene structure predictions within this context. Comparative analyses among members of jakobids and malawimonads further contribute to precise exon localization. We will present the nuclear gene content (proteins; non-coding RNAs including spliceosome components) across these flagellates, with an emphasis on spliceosomal introns. Malawimonad genomes range from 41 to 71 mbp in size, containing a remarkably large number (up to 147,000) of introns. In contrast, the nuclear genome size range of jakobids is only 20 to 55 mbp, with variable intron numbers (from only 1,592 up to ~100,000). Surprisingly, two malawimonads have the set of minor spliceosomal RNAs and the corresponding (probably ancestral) AT-AC introns, whereas there is no trace of these RNAs and introns in the examined jakobids. We will discuss an elevated potential for spliceosomal intron mobility across these protist flagellates.