THE NUCLEAR GENOME OF THE LICHEN-SYMBIONT MICROALGA *Trebouxia* sp. TR9 (Trebouxiophyceae, Chlorophyta). NEW ASSEMBLY AND ANNOTATION.

Marta Pérez-Rodrigo[1*]; Javier Montero-Pau[2]; Francisco Marco[1]; Patricia Moya[2]; Pedro Carrasco[1]; Eva Barreno[2]

[1] Universitat de València, Institut Universitari de Biotecnologia i Biomedicina (BIOTECMED), 46100 Burjassot, Spain; [2] Universitat de València, Inst. "Cavanilles" de Biodiversidad y Biología Evolutiva, Botánica, Fac. CC. Biológicas, Valencia, 46100 Burjassot, Spain; [*]E-mail: martaperezrodrigo@hotmail.com

Although genomic information concerning lichen-symbiont microalgae has increased, it is still very scarce compared to other organisms. However, this information is key to design adequate experimental strategies in ecophysiology and adaptation studies. So far, only the fully assembled chloroplastic and mitochondrial genomes of the lichen-symbiont microalgae *Trebouxia* sp. TR9, isolated from the lichen *Ramalina farinacea* (L.) Ach., were available. In this study, a first version of the nuclear genome was generated from 454 pyrosequencing and *Illumina* pair-end sequences; it displayed a high number of contigs, and its structural and functional annotation was limited. Therefore, the genome was sequenced again using *Oxford Nanopore*, which allowed longer sequences to improve scaffolding, closing, and gap filling. To improve the assembly and genome annotation, the transcriptome of *Trebouxia* sp. TR9 was sequenced and assembled *de novo*, resulting in the description of 92,389 transcripts by Trinity. The transcriptome was annotated using GO terms and the UniRef90 protein database, resulting in 52,010 annotated transcripts. This transcriptome represents the first collection of genes expressed in this species, and it constitutes an important genomic resource for ecophysiological studies. The sequences generated for the transcriptome were used for the reassembly of the genome. The result was a considerable improvement in the number of contigs, now 597, and their size, obtaining a definitive version of the genome of 59.7 Mb, which covers 100% of the estimated genome size and has a completeness of 96.7%. Furthermore, both the functional and structural annotation have been improved due to the integration of information from transcriptomes and the use of different annotation pipelines (BRAKER and MAKER-P). Regarding the structural annotation, the number of detected gene models is 16,391, and the functional annotation has been improved with a total of 7,068 different GO terms, 1,826 EC terms and 7,581 different gene annotations. Funding: PROMETEO/2017/039 (GVA).