Monodopsis and *Vischeria* Genomes Shed New Light on the Biology of Eustigmatophyte Algae

Hsiao-Pei Yang[1], Marius Wenzel[2], Duncan A. Hauser[1], Jessica M. Nelson[1], Xia Xu[1], Marek Eliáš[3], and Fay-Wei Li[1,4,*]

[1]Boyce Thompson Institute, Ithaca, New York, USA
[2]School of Biological Sciences, University of Aberdeen, Aberdeen, United Kingdom
[3]Department of Biology and Ecology, Faculty of Science, University of Ostrava, Ostrava, Czech Republic
[4]Plant Biology Section, Cornell University, USA

*Corresponding author: E-mail: fl329@cornell.edu.

Accepted: 9 October 2021

Abstract

Members of eustigmatophyte algae, especially *Nannochloropsis* and *Microchloropsis*, have been tapped for biofuel production owing to their exceptionally high lipid content. Although extensive genomic, transcriptomic, and synthetic biology toolkits have been made available for *Nannochloropsis* and *Microchloropsis*, very little is known about other eustigmatophytes. Here we present three near-chromosomal and gapless genome assemblies of *Monodopsis* strains C73 and C141 (60 Mb) and *Vischeria* strain C74 (106 Mb), which are the sister groups to *Nannochloropsis* and *Microchloropsis* in the order Eustigmatales. These genomes contain unusually high percentages of simple repeats, ranging from 12% to 21% of the total assembly size. Unlike *Nannochloropsis* and *Microchloropsis*, long interspersed nuclear element repeats are abundant in *Monodopsis* and *Vischeria* and might constitute the centromeric regions. We found that both mevalonate and nonmevalonate pathways for terpenoid biosynthesis are present in *Monodopsis* and *Vischeria*, which is different from *Nannochloropsis* and *Microchloropsis* that have only the latter. Our analysis further revealed extensive spliced leader *trans*-splicing in *Monodopsis* and *Vischeria* at 36–61% of genes. Altogether, the high-quality genomes of *Monodopsis* and *Vischeria* not only serve as the much-needed outgroups to advance *Nannochloropsis* and *Microchloropsis* research, but also shed new light on the biology and evolution of eustigmatophyte algae.

Key words: *Nannochloropsis*, spliced leader *trans*-splicing, simple sequence repeats, LINE, Stramenopiles.

Significance

Our current knowledge of eustigmatophytes mostly comes from the biofuel algae *Nannochloropsis* and *Microchloropsis*. Here we generated three high-quality genomes of *Monodopsis* and *Vischeria* that are sister to *Nannochloropsis* + *Microchloropsis*. We uncovered an extremely high prevalence of simple repeats in these genomes and found evidence of spliced leader *trans*-splicing. These new genomic resources will greatly facilitate future research to better understand the biology of eustigmatophytes, and to better capitalize on their translational potential.

Introduction

The diversity of algae is vast but largely unexplored. Despite their often inconspicuous nature, algae have played pivotal roles in Earth's biogeochemical cycles (de Vargas et al. 2015), and some might hold the key to sustainable bioenergy production (Radakovits et al. 2010; Jagadevan et al. 2018). Eustigmatophytes (Class Eustigmatophyceae), a lineage in Ochrophyta (Stramenopiles), are single-celled coccoid algae that can be found in freshwater, soil, and marine environments (Eliáš et al. 2017). The phylogeny and taxonomy of