



dr hab. Michał Szcześniak, prof. UAM
Adam Mickiewicz University in Poznań
Faculty of Biology
miszcz@amu.edu.pl

Poznań, July 14th, 2022

Review of Michał May's PhD thesis entitled
„PLASTICITY OF NUTRITION STRATEGIES IN MIXOTROPHIC
ORCHIDS AS AN INTERMEDIATE STAGE IN EVOLUTION OF
MYCOHETEROTROPHY”

The dissertation presented for review was supervised by prof. Marc-André Selosse and dr Marcin Jąkałski and is affiliated to the Faculty of Biology, Gdansk University. It has a classical layout, with a division into eleven chapters, such as Introduction, Methodology, Discussion or Bibliography, spanning 225 pages. Its core part represent five peer-reviewed publications as well as unpublished data, tackling diverse aspects of genetic basis of nutrition strategies in *Orchidaceae* family. The pieces of research are evaluated below.

1. Evaluation of the dissertation

1.1 A brief overview of published results

Article 1: *The complete chloroplast genome sequence of Platanthera chlorantha (Orchidaceae)*

This study provided a complete plastome sequence of *P. chlorantha*, creating a useful resource for phylogenetic and evolutionary studies. Such resources are especially important for the *Orchidoideae* tribe, as only 14 of its members' plastomes were available, with only one from the genus *Platanthera*. In a broader sense, the constructed plastome provides a necessary reference dataset describing a strictly autotrophic orchid, which serves as the foundation for further analysis of mixotrophic (MX) and autotrophic (AT) species.



Remarks:

- Please note that in the presented thesis, this article is dubbed „Article 1”, but appears as the second one in the subsequent chapters, which is a bit misleading. The same applies to the other published works, making the numbering of chapters completely off.
- Contribution of Michał May was substantial here and involved assembling, annotation and bioinformatics analysis of the plastidial genome as well as participating in writing the manuscript. Surprisingly, he appears as the second author here, while the first author, according to his statement, did not directly participate in the plastome assembly, which – I believe – is the main achievement of the work, nor did participate in writing the manuscript.

Article 2: *The complete chloroplast genome sequence of Dactylorhiza majalis (Rchb.) P.F. Hunt et Summerh. (Orchidaceae)*

In this paper, the authors assembled and analyzed the genome of *Dactylorhiza majalis* (Rchb.) P.F. Hunt et Summerh - the first fully sequenced plastome for its genus *Dactylorhiza*.

Contribution of Michał May was substantial here and involved assembling, annotation and bioinformatics analysis of the plastidial genome as well as participating in writing the manuscript.

Remarks:

- Both works regarding assembly of the plastidial genomes were published in the journal *Mitochondrial DNA Part B* – I understand the reason for that, nevertheless please clarify it.
- Both publications, though valuable, constitute little contribution to the broad goals of the dissertation project.

Article 3: *Three-year pot culture of Epipactis helleborine reveals autotrophic survival, without mycorrhizal networks, in a mixotrophic species*

The purpose of this study was to look into the mycorrhizal and nutritional status of *E. helleborine*, commonly considered as an mixotrophic species, that was growing in a pot without access to a mycorrhizal network with a photosynthetic carbon source. It sheds new light on the plasticity of mixotrophic nutrition and helps understanding its potential for autotrophy reversal.



A range of experimental and bioinformatics techniques were applied there, with a strong contribution of Michał May (such as isolation of the genetic material, fungal barcoding as well as „investigation process”), granting him the first authorship.

Article 4: *Thirteen new plastid genomes from mixotrophic and autotrophic species provide insights into heterotrophy evolution in Neottieae orchids*

This study looks at the effects of selective pressure on photosynthesis-related gene loss in the *Neottieae* tribe at various stages of evolution, from autotrophy to mycoheterotrophy. Surprisingly, photosynthesis-related genes in mixotrophic species did not experience selection relaxation when compared to autotrophic counterparts. In contrast, the authors found evidence of selection intensification for some plastid genes. Most relationships were resolved by phylogenetic analysis, but short branches at the base of the tree suggest an evolutionary radiation at the beginning of *Neottieae* history, which may be linked to the emergence of mixotrophy.

Remarks:

- Michał May is the sixth author there and his contribution involved collection of research material as well as data analysis. However, it is unclear which particular types of data processing and analysis were conducted by him, as for example Marcin Jąkałski declared data processing and analysis as well.
- Some genomes were assembled with Geneious, the others using CLC workbench – why different pieces of software were applied in the same study? With CLC, there were applied the following parameters: *discard short reads = yes* (limit 50 nt for MiSeq data, 25 nt for HiSeq data) – why different parameters for the two sequencers? Additionally, with BBDUK2 the parameter *minlength=10* was applied. Please comment on what is the lower limit for efficient assembly of organellar genomes.

Article 5: *The genomic impact of mycoheterotrophy in orchids*

This study compares the transcriptomic profiles of *E. aphyllum* and *N. nidus-avis*, two mycoheterotrophic (MH) orchids, at the organ level and between species in order to draw conclusions about gene loss and metabolic changes that plants go through during their transition to mycoheterotrophic nutrition.

The contribution of Michał May, was the following: collection of the research material, transcriptomic data analysis, debugging and testing the computational pipeline as well as



participating in preparation of the manuscript. Unfortunately, Michał May is the fourth author in this very interesting piece of research; nonetheless, his role in the research seems to be substantial.

Remarks:

- Why using BBMAP for expression estimation and not some of the more specialized software, such as RSEM?
- Hypergeometric test was part of gene set enrichment analysis: in what framework/software/programming language was it applied?

1.2 An overview of unpublished data

Unpublished work: *Practices for reliable hybrid de novo assembly of plant transcriptomes*

This study looks into several novel methods of transcriptomic assembly, with the goal of combining two cutting-edge sequencing platforms to overcome the limitations of *de novo* assembly processes. Before extending it to non-model research subjects, a comparison of multiple tools and approaches is performed to find the best settings for a model species (*A. thaliana*).

Here, Michał May is the first author and his contribution was bioinformatics data processing and preparation of the manuscript.

This chapter is formatted as a manuscript prepared for publication and perhaps represents the most interesting part of the thesis, at least when looking at Mr May's contribution in the light of scientific impact. Nonetheless I would like to raise a couple of questions which might help understanding some particular aspects of the study (due to their detailed nature, they do not need to be addressed during the public defense).

1. Read pre-processing: it is unclear whether it refers to Illumina or Oxford Nanopore data. What parameters were applied to identify adapters with LAST? Moreover, why was LAST applied there?
2. "Transcriptome assembly" chapter is followed by "Pre-assembly construction", which is quite confusing.
3. Only one plant species (*Arabidopsis thaliana*) was tested - the results obtained with another species might possibly differ substantially. If the work is supposed to be called „practices”, I would strongly suggest doing tests on some other species as well.
4. To assess the quality of assembled transcriptomes a number of tools were applied and they are enumerated (e.g. featureCounts or Bedtools) but it remains unclear what particular purposes they served.



5. To check the correctness and completeness of the assembled transcripts, the authors did several pieces of analysis, such as mapping against the genome. Would mapping against the transcriptome instead be a reasonable approach?
6. The authors applied filters to discard low-quality transcript candidates; how do they affect the potential to discover noncoding transcripts, such as long noncoding RNAs?
7. The authors write: „The highest percentage of genes covered in 100% of their length is present in IDP-denovo hybrid assembly based on velvet meta-assembly contig set (99.98%)” – was the assembly somehow perfect or maybe this statement should be rephrased?

1.3 A summary for the six pieces of research

Autotrophy (AT), the most common strategy used by plants, allows them to use solar energy to bind atmospheric or water-dissolved carbon dioxide into organic compounds. But in some plant species, an evolutionary adaptive regression resulted in the formation of mycoheterotrophy (MH), in which plants derive their entire organic carbon pool from a fungal partner found in a mycorrhizal relationship. This unusual strategy emerged as a result of a complex avalanche of evolutionary processes, and it coexists with multiple transient, intermediate stages required for its development, which is known as mixotrophy (MX).

The *Orchidaceae* are the largest family of flowering plants and one of the largest families in the *Plantae* kingdom. It contains nearly 28 000 described species. Because of their widespread occurrence and diversity, they serve as a useful model for general plant studies and evolutionary research. In the presented works, Mr May et al. conducted a comprehensive investigation of the genetic basis of nutrition strategies in the *Orchidaceae* family, including the autotrophy, mixotrophy, and mycoheterotrophy stages of trophic evolution. The dissertation is focused on how the adaptation to a specific trophism is reflected in changes in transcriptomes and plastidial genomes. It describes research into the basic state of plant autotrophy, plastome reduction and retention tendencies in the evolution of mixotrophic species, the possibility of adaptative reversal of mixotrophy, and changes in the genome that lead to the emergence of mycoheterotrophy – along with their consequences.

1.4 Evaluation: conclusions

According to the current regulations, including *art. 187 ust. 3 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce*, the review of PhD thesis is expected to answer the following three questions:



1. Does the dissertation present and document general knowledge of its author in the research area and discipline?

The presented dissertation in the discipline of „biological sciences” definitely shows broad knowledge and deep understanding of the studied phenomena by Michał May. First, there is an extended introductory part at the beginning of the thesis, including explanation of the biology behind evolution of nutrition traits in *Orchidaceae*. The methodology and theory behind it are nicely presented in a supplement (Chapter 10). Finally, the peer-reviewed and unpublished works themselves document well the author’s general knowledge.

2. Evaluation whether the dissertations shows its author’s ability to conduct scientific research on his own

I have no doubts, based on the authors’ declaration of contribution, that Michał May’s role was critical in at least four out of six pieces of research. In particular, the author gathered, unified, and optimized a set of tools for *in situ* procedures, laboratory sample handling, sequencing, and bioinformatic analysis pipelines. As a consequence, I am convinced that Mr May is a promising young scientist with a potential to do meaningful research.

3. Evaluation whether the presented dissertation constitutes an original solution to a scientific problem

The author broadened our knowledge in several directions, including the evolution of nutritional traits in plants and molecular biology aspects linked to it; also, a number of new plastid genomes were assembled and annotated. The observations helped to paint a picture of a plastidial genome that undergoes changes to effectively provide nutrition in response to changing conditions. Research on non-model organisms required development of custom solutions at different stages of research, from collection of samples to bioinformatics data analysis. Finally, Mr May contributes to development of best practices for *de novo* transcriptome assembly in plants. These and other findings presented in the dissertation clearly demonstrate solutions to scientific problems.

2. Final remarks

I find the presented dissertation valuable and inspiring. In general, the obtained results are of interest to the broad scientific community and my minor doubts and remarks raised above do not nullify my overall good impression. Besides its purely meritorical qualities, the conducted research was „a priceless personal learning opportunity and a lesson in non-linear exploratory approach”, as the author writes. I realize the studied subjects, except for *A. thaliana*, represent non-model organisms, hence Mr May had to face a number of challenges for this single reason, meaning the need to develop new methods or adjust the existing ones on a number of occasions. I do appreciate the information presented in Chapter 9.1: „Obstacles and opportunities in research process”, which clearly demonstrates how difficult it was to collect



the samples, to obtain pure DNA or RNA, and same with some other wet lab experiments and *in silico* procedures. This makes me appreciate the work done by Mr May even more and conclude that the dissertation meets all the requirements for the doctoral thesis. I, therefore, strongly recommend proceeding to the following steps of the doctoral defense as well as conclude that the dissertation should be awarded appropriately.

Michał Szczerba

ul. Uniwersytetu Poznańskiego 6, Collegium Biologicum, 61-614 Poznań
NIP 777 00 06 350, REGON 000001293
tel. 61 829 5945

www.biologia.amu.edu.pl