



To the dean of the
Faculty of Science
Prof. Dr. Ing. Hana Santuckova
University of South Bohemia
in Ceske Budejovice

www.igb-berlin.de

Leibniz-Institute of
Freshwater Ecology and
Inland Fisheries

im Forschungsverbund
Berlin e.V.

Date: 3rd of January 2023

Assessment of the doctoral thesis of Mr. **Vinicius Silva Kavagutti** with the topic

Microbial and viral genomic diversity in Freshwater environments

Topical relevance

The present thesis has a high topical relevance as it addresses several of the most challenging topics in plankton ecology and limnology.

Freshwater microbial communities are complex and highly dynamic both in space and time, and often show distinct, recurrent patterns controlled by biotic and abiotic environmental factors. Yet, current studies often miss the accurate spatial and temporal scales to reliably analyze patterns of recurrence of individual pelagic members including viruses, prokaryotes, eukaryotes. In particular, limitation in discriminating and quantifying each plankton member poses a great problem in the field.

Also, our limited ability to achieve cultures of environmentally abundant microbes from various aquatic ecosystems still hinders our capabilities to disentangling complex planktonic food webs shaped by numerous ecological interactions of microbial key players (viruses, pro- and eukaryotes). In particular, viruses - the numerically most abundant actors in planktonic food webs – are relatively few studies, especially in freshwaters. Consequently, little studies have addressed their identities and ecology.

These major knowledge gaps in aquatic microbial ecology have increasingly gained attention with the development of various OMICS tools. Recently, major advance in bioinformatics tools have enabled the recovery of even complete genomes from ultra-deep metagenomic sequencing of planktonic samples. These techniques allow to simultaneously identify the microbial hosts and their associated virus. Moreover, they allow to identify specific functional features not only of the pro- and eukaryotic hosts but also of their associated phages/viruses. Thereby, it is possible to identify the ecological role of viruses, i.e. to evaluate whether they can also have positive effects for their host, e.g. via horizontal gene transfer (HGT) of functional genes impacting both physiology and evolution of their hosts. This study helps to revisit the ecological role of viruses in planktonic food webs and their importance for biogeochemical processes and plankton dynamics both in space and time. As such the study holds a great potential to close some of our current knowledge gaps in aquatic microbial ecology and to change our concepts on planktonic food webs and their role in global biochemical cycles.

Prof. Dr. Hans-Peter Grossart

Head of the department "Plankton and
Microbial Ecology" at IGB

Head of the research group
"Aquatic Microbial Ecology"

Plankton and Microbial Ecology

Zur alten Fischerhütte 2
16775 Stechlin – Neuglobsow
Germany

Tel: +49 (033082) 699 91

Fax: +49 (033082) 699 17

hgrossart@igb-berlin.de

Institute for Biochemistry and Biology
Potsdam Univ., Am Neuen Palais 10,
D-14469 Potsdam

The present PhD thesis has addressed the microbial and viral diversity in a well-studied site, i.e. the Římov reservoir, that is part of a long-term monitoring program (for several decades) of the Biology Centre CAS, Institute of Hydrobiology in Budweis. The study has used long-term low-frequency seasonal sampling as well as short-term high-frequency sampling strategies to illuminate the microbiological dynamics at this site in a rather unbiased approach. By using metagenomics to obtain genomes of both microbial hosts and their viruses, the PhD candidate was able to reveal their diversity, seasonality, distribution, and interactions at the different time scales of observation. Thereby, Vinicius Silva Kavagutti could identify an extremely high microbial diversity which to a large extent has not been cultured yet and only limited knowledge on their ecological role in planktonic food webs existed before. The most impressive result of the study is that the freshwater planktonic community is extremely diverse - also at short temporal scales, and their genome analyses still reveals a lot of ecological surprises. As such the study is able to change our current perception on the role of microbes in freshwater planktonic food webs, in particular that of virus. I am convinced that the present PhD thesis leaves its foot print in limnology and, especially in aquatic microbial ecology.

Content of the thesis

The thesis is written in English and consists of a general introduction, summaries and conclusions of the three studies performed in this PhD thesis, as well as three main chapters (with the 3 publications, meanwhile all accepted in highly ranked, peer-reviewed journals, i.e. 2 in Microbiome and 1 in mSystems), with in total 106 pages including a bibliography and lists of tables and figures.

Chapter 1 reports on a metagenomic study (published in “Microbiome”) to illuminate and better understand the ecology of bacteria and viruses. By using metagenomic data sequenced from 2 temperate freshwater habitats (Římov Reservoir and Jiříčka pond), genomes of phages (Caudovirales) and of their hosts could be reconstructed. Overall, the study revealed 2034 complete genomes of phages infecting abundant freshwater prokaryotes, some of which are predicted to infect freshwater Actinobacteria, Betaproteobacteriales, Alphaproteobacteria, Bacteroidetes, Chloroflexi, and other prokaryotic groups for which no phages have previously been described. The study also recovered 444 actinobacterial MAGs (metagenome assembled genomes), the majority assigned to ‘Ca. Nanopelagicales’ (n = 280), Acidimicrobia (n = 114), and Microbacteriaceae (n = 37). Interestingly, the respective abundance patterns coupled with phylogenetic analysis indicates that even closely related genomes can have different environmental preferences. In particular, they can show very different abundance peaks in the epi or hypolimnion of the studied systems. In addition, the study used abundant freshwater Actinobacteria and their phages as model systems. Their genomic analyses reveal that phage genome abundances in deep water bodies mirror the abundance of their hosts and thus reflect classical patterns in thermal cycles of the water column. High phage and host abundances in the epilimnion are transitory, whereas both tend to persist at lower abundances in the hypolimnion. Also, genomic phage analysis shows a widespread prevalence of an AB toxin gene predicted to inhibit eukaryotic actin polymerization. This notion indicates phages may carry genes helping to protect the host from eukaryotic predation acting as a quasi “Trojan Horse”. A remarkably high content of anti-oxidative genes in phage genomes suggest that they improve survival of infected hosts when ingested they are ingested by eukaryotic grazers (i.e. phage-mediated ROS defense). Both mechanisms (Trojan Horse and Phage mediated ROS defense) imply that eukaryotic grazing pressure acts as a strong selective force with visible effects in the phage genomes.

Chapter 2 consists of a paper (accepted in “Microbiome”). This study represents a high-resolution ecological dissection of a spring bloom in the Římov reservoir. It combines high-throughput sequencing with classical microbiological techniques to address the short-term (daily sampling for 37 days at 2 different depths) dynamics of diversity patterns across all domains of life (eukaryotes, prokaryotes, and viruses). The study comprises 57 metagenomic samples of 3 different size fractions (5 μm , 0.8 μm , and 0.22 μm) and thus captures the entire cascade of spring bloom events. Also, different filtration methods were compared.

Reconstruction of in total 2214 bacterial MAGs indicated the dominance of several bacterial phyla, i.e. Bacteroidota (n=690), Proteobacteriota (n=677), and Actinobacteriota (n=453). Thereby, the multiple distributional dynamics for several bacterial groups (e.g., ‘Ca. Fonsibacter’, ‘Ca. Planktophila’, and Limnohabitans) with progressively increasing stratification could be revealed. Interestingly, even very closely related genomes can show quite different niche preferences and thus distribution patterns. The different observed bacterial distribution patterns were assigned to 4 bacterial strategies, primarily based on their temperature preferences and occurrence in the epi- vs. hypolimnion. Combining MAG-derived abundances with CARD-FISH analyses revealed similar in-situ doubling time estimates (ca. 2 days) for all dominant genome-streamlined microbial lineages (‘Ca. Nanopelagicus’, ‘Ca. Methylopumilus’ and ‘Ca Fonsibacter’).

Also, 679 complete bacteriophage genomes could be reconstructed, whereby only 129 had been described before in the group’s publications. Thus, they may be considered persistent phages in the Římov reservoir. The study presents 35 genomes predicted to infect Actinobacteriota and three others Cyanobacteria. Interestingly, this PhD describes a new host gene (RNA polymerase sigma factor, TIGR02997) seemingly prevalent in freshwater cyanophages and. Thus, in future studies it can serve as a marker for cyanophage predictions from metagenomic sequences. In this study, 3309 additional Nucleo-Cytoplasmic Large DNA Viruses (NCLDV, phylum Nucleocytoviricota, or giant viruses) contigs were recovered, of which most infect Haptophytes like *Chrysochromulina* (99 contigs) and *Phaeocystis* (163 contigs). This confirms the ecological role of viruses in top-down control of bacteria and unicellular eukaryotes (the viral shunt) and thereby releasing substantial amounts organic and inorganic nutrients. The study nicely integrates these findings in the current PEG model and also indicates uncertainties in the role of parasitic perkinsids vs viruses in phytoplankton bloom. Interestingly, the study suggests that the NCLDVs, which are more prevalent throughout the water column, may be more responsible for bloom collapse than the perkinsids that also prefer deeper and colder water layers. Previous studies confirm this key ecological role of NCLDVs in planktonic food webs. From these findings and literature data, the study concludes a combined effect of NCLDV, zooplankton grazing as top-down controls, and nutrient limitation as a bottom-up control for the recurrent collapse of the phytoplankton spring bloom.

Due to the high frequency of sampling the first metagenomic-assembled genomes of freshwater protists (a diatom and a haptophyte) could be generated. Further, microscopic identification revealed *Cyclotella* as a dominant centric diatom, whereas phylogenomic analyses assign it to *Thalassiosira*, which has been primarily described as a marine species. Conclusively, there is a high need to link microscopic identification and genome sequence methods for a more proper species identification. The study also confirms the importance of cryptophytes in the spring bloom, in particular of aplastidic cryptophytes, which have been very recently identified as key members of the eukaryotic community in freshwaters. The simultaneous application of CARD-FISH for the spring bloom analyses supports the dominance of a yet hidden and abundant aplastidic cryptophyte during the spring bloom in the Římov reservoir.

Chapter 3 reports on an extensive search of hundreds of metagenomes and metatranscriptomes, which led to the discovery of 399 rhodopsins with unusual amino acid substitutions and novel motifs, which were named Alt-rhodopsins (AltRs). The discovered rhodopsin diversity allowed to recreate large-scale, well-supported phylogenies and provides a reliable basis for heliorhodopsin genesis and evolution. Novel rhodopsin sequence variations could be uncovered by analyzing different databases with numerous genomic, metagenomic, and metatranscriptomic data sets from various studies. The presented downstream analysis resulted in several new and undescribed variants of the residue 82 (Arg82), comprising multiple rhodopsin sequence variants with different replacements for the residue 82 (e.g. H, K, Q, A, P, S, Y, E, and M).

Out of the retrieved 399 sequences, only 121 have been previously described as channelrhodopsins, xenorhodopsins, and potassium pumps, which all represent Type-1 rhodopsins. The study shows 278 novel AltRs encompass Type-1 and HeRs from several phylogenetic clades such as eukaryotes (n=30), giant viruses (n=34), bacterial (n=25), and archaeal (n=4). Yet, 185 sequences remained unclassified. Based on the retrieved large collection of rhodopsins a reliable phylogenetic reconstruction of the rhodopsins could be performed. It revealed a widespread convergent evolution of AltRs across all domains of life. Importantly, the specific tree topology suggests that HeRs originate from eukaryotes and then occurred in prokaryotes, contrary to what is more commonly observed.

The thesis provides a comprehensive summary and conclusions for each chapter highlighting the main findings of this study and puts them in context to previous references and the current state of knowledge. This part of the PhD thesis nicely demonstrates that the obtained results have far-reaching implications for key concepts of microbial ecology in freshwaters. In particular, the findings of numerous novel prokaryotes, viruses and functional genes help to greatly improve our knowledge in temporal and spatial dynamics of freshwater microorganisms. The many new and exciting findings of this study certainly help to redefine current concepts and theories in microbial ecology and understand the role of microbial diversity and functions for the biochemical cycle in freshwaters.

Evaluation

Mr. **Vinicius Silva Kavagutti** has made a very significant scientific contribution to aquatic microbial ecology, and more specifically in viral and prokaryotic dynamics both in the long- and short-term. His genomic analyses of microbial diversity from viruses, archaea, bacteria to eukaryotic microorganisms form a solid fundament for future phylogenetic and functional analyses based on metagenome and metatranscriptome datasets. He also sets his exciting results in the context of the PEG model and thereby identifies and closes current knowledge gaps. The work is based on a massive amount of sequence data and thus allows to relate identified patterns in microbial dynamics to the ecological role of specific microbial groups. His work without any doubt represents a big step forward in aquatic microbial ecology, in particular in limnology.

Mr. Vinicius Silva Kavagutti has applied numerous microbiological, genomics and statistical tools to decipher temporal and spatial dynamics of microbial communities in freshwaters related to environmental changes which may also help to evaluate consequences of current and future climate change. He undoubtedly worked very hard and is first author of 3 journal papers (2 published) and 1 accepted. At the same time, he is coauthor of four additional international peer-reviewed papers. The thesis is well written and doesn't contain many language errors and

hence I have the impression that the thesis was very carefully written. I only have some minor comments to the thesis (which I have added in an attached pdf-file of the PhD thesis). Of course, there could be more analyses, e.g. metatranscriptome analyses, to better nail down the precise functions of specific microbial groups. Also, in my opinion, a summary graph and may be a more comprehensive conclusion chapter could be nice additions to better summarize and highlight the obtained results. This would make it easier for the reader to relate the findings of this study to a broader, more global picture. However, this comment shouldn't be seen as a negative remark since overall the PhD thesis leaves an excellent impression.

Consequently, I evaluate the submitted doctoral thesis of Mr. Vinicius Silva Kavagutti with **excellent** and **recommend that the above-mentioned research work should be accepted.**

Yours sincerely,

A handwritten signature in blue ink, appearing to read 'H.P. Grossart', with a long horizontal stroke extending to the right.

(Prof. Dr. Hans-Peter Grossart)