University of South Bohemia in České Budějovice

Faculty of Science

Department of Molecular Biology and Genetics

Laboratory of Environmental Genomics

Molecular mechanisms of predation of planktonic protist *Hemistasia phaeocysticola*

Master thesis

Bc. Tomáš Knobloch

Supervisor: Mgr. Aleš Horák, Ph.D.

Supervisor – specialist: Mgr. Valéria Juricová

České Budějovice, 2021

Knobloch T., 2021. Molecular mechanisms of predation of planktonic protist *Hemistasia phaeocysticola*. Mgr. thesis, in English – 42 p., Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic.

Annotation

In this thesis growth curves under specific conditions, fluorescence microscopy, and transcriptomic analyses were used to investigate whether the marine protist *Hemistasia phaeocysticola* is able to prey on bacteria and the determination of its molecular mechanisms of predation.

Affirmation

I hereby declare that I have worked on this Master thesis independently and used only the source listed in the bibliography.

I hereby declare that, in accordance with Article 47b of Act No. 111/1998 in the valid wording, I agree with the publication of my Master thesis, in full form to be kept in the Faculty of Science archive, in electronic form in a publicly accessible part of STAG database operated by the University of South Bohemia in České Budějovice accessible through its web pages.

Further, I agree to the electronic publication of the comments of my supervisor and thesis opponents and the record of the proceedings and results of the thesis defense in accordance with aforementioned Act No. 111/1998. I also agree to the comparison of the text with the Thesis.cz thesis database operated by the National Registry of University Theses and a plagiarism detection system.

Date..................................

Signature............................

Acknowledgment

I would like to express my gratitude to my supervisor Mgr. Aleš Horák, Ph.D. for allowing me to work on this Master thesis in his laboratory, his expertise and advices throughout the whole project. Next, I would like to thank my supervisor – specialist Mgr. Valéria Juricová for her help with designing experiments and consulting all difficulties. Further, I am thankful for advices in times of need from all members of the Laboratory of Environmental Genomics and Laboratory of Molecular Biology of Protist. I am also thankful to Prof. RNDr. Julius Lukeš, CSc. For letting me use equipment in his Laboratory of Molecular Biology of Protist for several experiments for this thesis. And my greatest gratitude goes to my family and friends for endless support during my studies and my whole life.

Table of content

1. Introduction

1.1. Ocean ecosystems

The ocean covers 71% of the surface of the Earth and amounts to about 97% of all water on our planet. It can be vertically divided into zones based on the depth and light abundance. The Photic zone ranges from the surface to 200 m of depth. There is enough light to enable photosynthesis and therefore it is a zone with the most biodiversity. Deeper zones mostly depend on sinking and dissolved nutrients or hydrothermal vents as sources of energy. The second region is Mesophelagic ranging from 200 m to thermocline boundary of 12˚C generally at 700 m to 1 000 m of depth. The Bathypelagic zone lies between 12˚C and 4˚C thermoclines up to 4 000 m deep, followed by sparsely researched Abyssalpelagic zone and Hadalpelagic region of Mariana Trench. (Edgcomb, 2016)

The global ocean has a major impact also on land. The ocean temperature determines climate and wind patterns and is the main component of the water cycle. (Schmitt, 1995b) It was the birthplace of life on Earth around 3 billion years ago. It is the biggest ecosystem on Earth in regard to both area and volume and hosts the largest biomass, most of which is composed of microorganisms - plankton. The prokaryotic and eukaryotic photosynthetic plankton rivals the land plants in primary production. Compared to these, plankton´s diversity, life, and nutrients cycles are still much less understood. The importance of viruses, bacteria, cyanobacteria, algae, and other photosynthetic protists is well ingrained and their interaction with heterotrophic protists forming intertwined food webs is slowly uncovered. (de Vargas et al., 2015), (Worden et al., 2015), (Pernice et al., 2014)

In the past, the carbon cycle in marine plankton was thought to be mainly through the bacterial loop – from autotrophic protists to heterotrophic bacteria to heterotrophic protists. (Skovgaard, 2014) Bacteria are ubiquitous in marine ecosystems. (Lalli & Parsons, 1997) They are unchallenged in terms of diversity, abundance, and metabolic activity and are major players in carbon flow due to their uptake of organic matter, which is able to change the overall carbon cycle in oceans and, therefore, globally. (Azam & Malfatti, 2007) Heterotrophic protists feeding on bacteria were thought to be the main factor for keeping in check the high reproduction rate of bacteria in the oceans and getting most of the nutrients. (Sherr & Sherr, 1994) The ways for more accurate measurements are ever-evolving, as shown by *in situ* approaches instead of traditional

shipboard incubations, which can lead to a high degree of bias introduced by inevitable processes such as ambient light during transport of seawater to the surface and initiation of incubation, pressure changes and other physical and chemical parameters. (Pachiadaki et al., 2016) An article showed protists´ grazing on prokaryotes in bathypelagic deep-waters by fluorescently-labeled prey analogues, where protists were able to remove 3,79% (±1,72%) to 31,14% (± 8,24%) of the prokaryote stock. This implies the importance of protist grazing even in bathypelagic depths. (Rocke et al., 2015) Although the concept of bacterial loop functions to some extent, there was a gradual change in perspective of plankton ecology during recent decades. Presently, the bacterial loop is viewed only as a part of the ever-expanding intricate nutrient cycle in the marine ecosystem. It was recognized, that a significant number of planktonic protists cannot be easily categorized into heterotroph and autotroph categories. An increasing number of protist groups exhibits various degrees of mixotrophy – an ability to utilize both autotrophic and heterotrophic pathways. This further complicates accurate assessments of oceanic nutrient cycles and marine planktonic food webs. (Worden et al., 2015), (Stoecker & Pierson, 2019)

In the past, heterotrophic protistian grazers were mainly considered to be grazing on a much smaller prey, like bacteria. (Sheldon et al., 1977) During the last few decades, more attention was given to heterotrophic protists, which gain nutrients by preying on other protists. (Skovgaard, 2014, p. 51) Some of the dominant microscopic eukaryotic predators, like dinoflagellates, were found to be preying on organisms of a similar size. (Hansen et al., 1994) Moreover, a number of heterotrophic protists are preying on a bigger prey than their own cells, which can be considered a form of parasitism, referred to as parasitoidism. Parasitoid is an organism that lives in close association with its host at its expanse, resulting in host´s death. (Skovgaard, 2014) There are protistan classes, for example Protalveolata and Apicomplexa, which are comprised primary of parasites. Parasitic and parasitoidic protists turn up to be a major cause of mortality for other protistan taxa such as dinoflagellates and marine alveolata group II. (Edgcomb, 2016) Not much is currently known about these feeding strategies, as they were just recently considered important and at high enough scale to make a difference in the food webs of marine plankton. (Skovgaard, 2014)

1.2. Heterotrophic protists

Planktonic heterotrophic flagellates are significant members of marine ecosystems. They are single-cell organisms who possess one-to-many flagella for motility and are commonly less than 20 μm long. (Logares et al., 2012) When grouped by size, planktonic heterotrophic flagellates can be described in their roles in the marine nutrient cycle. Pico-heterotrophic flagellates range from 1 μm to 5 μm and are mostly marine bacteriovores with a key role as intermediates for energy transfer between bacteria and larger protists. They are common prey to other organisms, for example metazoans and ciliates. (Massana, 2011b) Micro-heterotrophic flagellates, 5 μm to 20 μm long, including bodonids, rhizarians, and ciliates, effectively control prokaryotic biomass. (Calbet & Landry, 2004b) Heterotrophic protists were found in all sampled depths of the oceans with decreasing abundance ranging from an average 72 ± 19 cell ml^{-1} in mesopelagic waters to 32 ± 3 cell ml^{-1} in bathypelagic waters to 11 ± 1 cell ml^{-1} in the deepest measured layers. (Pernice et al., 2014)

Tara Ocean expedition approached the research and the sampling of oceanic plankton differently from other studies, which were more region focused. They collected samples from 334 zones across tropical and temperate oceans and analyzed V9 barcodes, which are short fragments of 18S ribosomal RNA gene - phylogenetically informative for its variability, of eukaryotes ranging from single-cell organisms to small animals a few millimeters in size. From approximately 150 000 operational taxonomic units, one-third could not be assigned to any known eukaryotic group. (de Vargas et al., 2015), (Lukeš et al., 2015)

1.3. Diplonemids

From the findings of Tara Ocean expedition, diplonemids stood out with their unexpectedly high diversity and abundance. (Flegontova et al., 2016) Diplonemida order mainly consists of unicellular biflagellated heterotrophic protists. They belong to the phylum Euglenozoa together with many important species such as photosynthetic algae *Euglena* and pathogens to livestock and humans for example *Trypanosoma* and *Leishmania*. (Vesteg et al., 2019b) Diplonemids have cosmopolitan presence in shallow littoral sediments, deep aphotic pelagic waters, abyssopelagic zone, hydrothermal vents, and even in deep fresh water lakes. (Flegontova et al., 2016), (Mukherjee et al., 2019) Phylogeny based on 18S rRNA categorized diplonemids as a monophyletic group divided into four

lineages. Diplonemidae clade contains dozens described species with described morphology and several cultivatable specimens. *Hemistasia phaeocysticola*, as an organism of interest for this thesis, is a member of Hemistasiidae clade. Deap-sea pelagic diplonemids (DSPD I – recently categorized as Eupelagonemidae and DSPD II) clades are responsible for the majority of diversity and abundance in world´s oceans, coming from mainly environmental sequences and single cell isolates. The described diplonemids exhibit similar morphological and ultrastructural features. The cells are units to tens of μm long exhibiting elliptical, pouchlike, and highly variable shapes with great plasticity of movement owing to naked plasma membrane supported by tightly packed microtubular frame. There is an apical papillum at the apical end of the cell nearby of which is the feeding apparatus and the pouch from which emerge two flagella. (Tashyreva et al., 2018a)

Similarly to kinetoplastids, diplonemids have a single peripheral highly reticulated mitochondrion containing vast mitochondrial DNA made up of thousands of concentrated circular molecules. Mitochondrial genomes of diplonemids consist of dozens of distinct circular chromosomes which can greatly differ in sizes. One described outlier is *Diplonema papillatum* with 260 Mbp of DNA in its mitochondrion making it the highest amount of DNA in a single organelle reported to date. (Lukeš et al., 2018a) Other stand out features are extensive trans-splicing and numerous forms of RNA editing. (Valach et al., 2017)

The feeding patterns of diplonemids are not yet fully uncovered. So far, evidences were presented describing diplonemids preying on variety of prokaryotes and eukaryotes, such as algae, to varying degrees species to species. (Lukeš et al., 2015) Parasitismus of Planktonic diatoms was also described in one diplonemid species, *Rhynchopus coscinodiscivorus*. (Schnepf, 1994) Phototrophy and mixotrophy are generally excluded as probable feeding strategies of diplonemids due to large abundance of diplonemids in deep ocean and phylogenetic relations. (Flegontova et al., 2016)
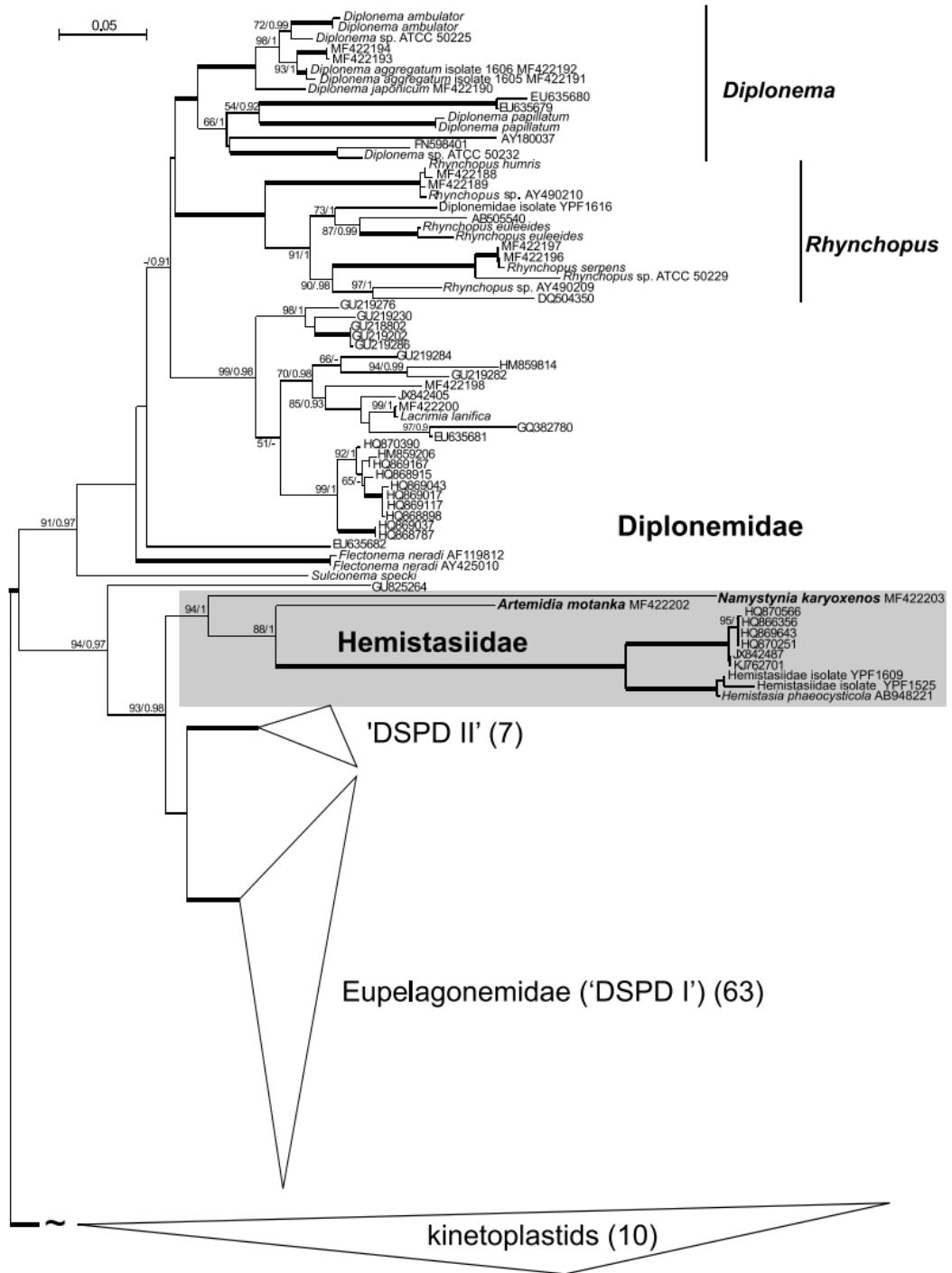
Fig. 1.: Phylogenetic positions of current diplonemids by maximum likelihood analysis of 18S rDNA. Highlited are newly described taxa. Thick branches represent absolute support. (Prokopchuk et al., 2019)

1.4. Hemistasia phaeocysticola

Hemistasia phaeocysticola is one of the model organisms of diplonemids, because it can be easily cultivated in seawater based medium. (Lukeš et al., 2015) It belongs to the family Hemistasiidae, closely related to the family Diplonemidae and to Kinetoplastids. (Prokopchuk et al., 2019) *H. phaeocysticola* is polykinetoplastic biflagellate protist with smooth cell surface. The cell size and shape vary largely ranging from approximately 10 µm to 25 µm long and 3 to 7 µm thick. Hungry cells are prolonged and active (Figure 2); well fed cells are pear-shaped or even near spheres (Figure 3). It has small nucleus near the dorsally located flagellar pocket at the anterior region of the cell; cytostome situated on the ventral sub-anterior side and single large digestive vacuole occupying most of the posterior region. (Elbrächter et al., 1996) *H. paheocysticola* has branched mitochondrion with lamellar cristae containing dense clusters of DNA. (Prokopchuk et al., 2019) The division of the cell happens by binary fission before which most of the cells are attached with their rostrum to the substrate. *H. phaeocysticola* swims fast in strait line, in spiral and can quickly change directions, but shows only scarce movement when in a rich-in-nutrients environment. It takes nutrients in small portions from multitude of decaying organisms or attacks living ones. Some of the described prey includes diatoms for example *Thalassiosira rotula* and *Coscinodiscus granii*, dinoflagellates *Gonyaulax polyedra*, haptophyte *Phaeocystis globusa* and even dead *H. phaeocysticola* cells. (Elbrächter et al., 1996)

Fig. 2.: Hemistasia phaeocysticola with highly active prolonged phenotype in an environment containing low concentrations of nutrients or not fed on prey.
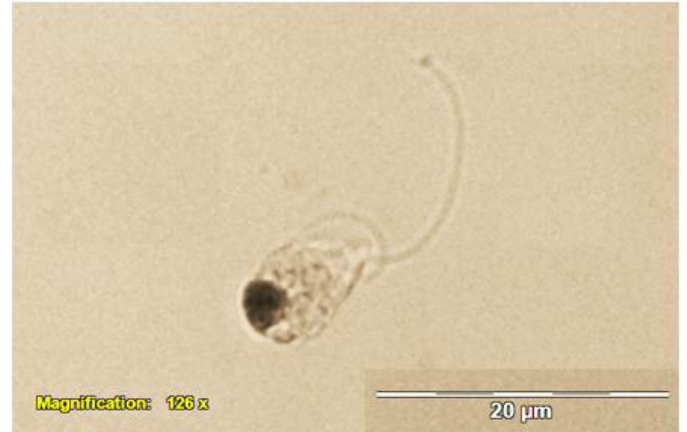
Fig. 3.: Hemistasia phaeocysticola with pear-shaped phenotype in an environment containing high concentrations of nutrients or well fed on prey.

2. Aim of the thesis

The aim of this thesis was to investigate molecular mechanisms of predation of heterotrophic flagellate *Hemistasia phaeocysticola*. For this purpose we measured growth curves during various growth conditions (medium, seawater medium, and preying on bacteria), fluorescence microscopy and bioinformatic analysis.

Working hypotheses:

- *H. phaeocysticola* is a generalist heterotrophic flagellate capable of feeding on bacteria
- As such, it doesn´t discriminate between different sources of energy

The aims based on hypotheses:

- To measure the ability of *H. phaeocysticola* to grow on bacteria
- To analyze differentially expressed genes associated with feeding on bacteria and compare transcriptomic profiles between different trophic modes

3. Materials and methods

3.1. Determination of bacterial prey for *Hemistasia phaeocysticola*

The bacteria from the mixed culture of *Hemistasia phaeocysticola* were plated onto Petri dish containing Marine broth (Difco[TM]) mixed with agarose (10 g/l). Plates were incubated at 32˚C for 16 hours. Grown cultures were used as a template for PCR (Table 1, 2).

Tab. 1.: The composition of PCR to reveal the genus of the bacterial prey of *Hemistasia phaeocysticola*.

Solution	Volume
Final Volume	25 µl
OneTaq 2x Master Mix	12,5 µl
16S EUB forward primer 5´- GCTTAACACATGCAAG - 3´	0,5 µl
16S EUB reverse primer 5´- CATTGTAGCACGTGT - 3´	0,5 µl
Bacterial colony	part
MilliQ H$_2$O	11,5 µl

Tab. 2.: The PCR conditions for revealing the genus of the bacterial prey of *Hemistasia phaeocysticola.*

	Temperature	Time	Cycles
Preheating	95 °C	3 min	1
Denaturation	95 °C	40 sec	
Annealing	55 °C	40 sec	34
Extension	72 °C	1 min	
Final Extension	72 °C	10 in	1

3.1.1. *Escherichia coli* transformation for determination of bacteria genus serving as prey for *Hemistasia phaeocysticola*

The PCR product was purified using PureLink[TM] PCR Purification kit (Invitrogen). 4µl of purified PCR product were mixed with 0,5 µl of TOPO vector, 0,5 µl of Salt Solution and 1 µl of deionised H$_2$O and the mixture was incubated at room temperature for 15 min. 3 µl of mixture were added to 50 µl of Chemically competent *E. coli* XL-1 Blue and were

left to incubate on ice for 30 min. The solution was subjected to a heat shock of 42 °C for 30 sec and immediately transferred on ice. 250 µl of SOC medium (Thermo Fisher Scientific) were added to the mixture and chemically competent *E. coli* XL-1 Blue were allowed to recover at 37 °C subjected to 200 rpm shaking for 1 h. Transformed *E. coli* were plated onto ampicillin agar plate (Table 3) subjected to prior treatment of a mixture containing 40 µl of X-gal (40 mg/ml) (Thermo Fisher Scientific), 40 µl of deionised H_2O and 4 µl of IPTG (200 mg/ml) (Thermo Fisher Scientific) and were left at 37 ˚C over night.

Tab. 3.: The composition of agar plates.

Solution	Amount [800 ml]
Yeast extract	4 g
NaCl	8 g
Tryptone	8 g
Agar bacteriological	8 g
Ampicilline	3,2 ml
Milli Q H_2O	800 ml

3.1.2. Determining of bacterial a prey for *Hemistasia phaeocysticola*

Colonies of successfully transformed *E. coli* were used as a template for PCR (Table 4, 5). The obtained PCR product was subjected to agarose gel electrophoresis (Figure 4) to ensure colonies contained the desired product and was then purified using PureLink^{TM} PCR Purification kit (Invitrogen). The purified PCR product was sent for sequencing using Mix2Seq Kit (Eurofins).

Tab. 4.: The composition of PCR for sequencing of bacterial prey for *Hemistasia phaeocysticola*.

Solution	Volume
Final Volume	50 µl
OneTaq 2x Master Mix	25 µl
M13 forward primer 5´ - GTAAAACGACGGCCAG – 3´	0,5 µl
M13 reverse primer 5´ - CAGGAAACAGCTATGAC – 3´	0,5 µl
Bacterial colony	part
MilliQ H$_2$O	24 µl

Tab. 5.: The PCR conditions for sequencing of bacterial prey for *Hemistasia phaeocysticola.*

	Temperature	Time	Cycles
Preheating	95 °C	3 min	1
Denaturation	95 °C	40 sec	
Annealing	50 °C	40 sec	34
Extension	72 °C	1,5 min	
Final Extension	72 °C	10 min	1

3.2. Cell cultures

3.2.1. *Hemistasia phaeocysticola*

The culture of *Hemistasia phaeocysticola* was maintained in Hemi medium (Table 6) at 15˚C without light. The culture was subcultured once a week by 1:3 dilutions (1 part culture, 3 parts of medium).

Tab. 6.: The composition of Hemi medium.

solution	Amount
Sea salt (sera marin)	36 g/l
Trypton	10 mg/l
Yeast extract	5 mg/l
NaCl	10 mg/l
Horse serum	10 ml/l

3.2.2. *Paracoccus sp.*

The culture of *Paracoccus sp.* was maintained in liquid Marine Broth medium (Difco[TM]) at 15˚C with no light and was subcultured once a week by taking a small amount of grown culture and transferred into the desired amount of Marine Broth medium (Difco[TM]).

3.3. Growth curve of *Hemistasia phaeocysticola* preying on live *Paracoccus* sp.

To determine whether *Hemistasia phaeocysticola* can survive and thrive on *Paracoccus* sp. as the only source of nutrients a growth curve was prepared. *H. phaeocysticola* cells were maintained in a lower nutrient medium (1/4 of nutrient concentration of Hemi medium [Table 6]) for a minimum of 4 weeks prior to the growth curve analysis to allow the cells to adjust to a lower amount of available nutrients. $2 \cdot 10^3$ cells of *H. phaeocysticola* maintained in a low nutrient medium were harvested by centrifugation at 3000 g for 10 min for each of three separate growth curve cultures, one negative and one positive control cultures. Harvested cells were transferred into 9 ml of artificial seawater (36 g/l of sea salt [sera marin]) with no additional nutrients. 1 ml of grown *Paracoccus* sp. culture per each of three separate growth curve cultures was harvested by centrifugation at 4000 g for 5 min. The pellet was resuspended in 1 ml of artificial seawater (36 g/l of sea salt [sera marin]) and the solution was transferred into growth curve culture resulting in 10 ml volume of each of the growth culture. The negative control growth curve culture contained only $2 \cdot 10^3$ cells of *H. phaeocysticola* in 10 ml of artificial seawater (36 g/l of sea salt [sera marin]) without any additional nutrients. The positive control growth curve culture contained only $2 \cdot 10^3$ cells of *H. phaeocysticola* in 10 ml of Hemi medium (Table 6). The cultures were kept at 15˚C

for nine days with a sampling period of 24 hours. The counting of cells was done using Neubauer improved Counting Chamber (Marienfield, Germany).

3.4. Fluorescence in-sittu hybridization with universal bacterial probe (Eub338)

Roughly 10^6 of *Hemistasia phaeocysticola* cells from the mixed culture of *H. phaeocysticola* with *Paracoccus* sp. per each slide were harvested by centrifugation at 3000 g for 10 min. The pellet was resuspended in 250 µl of artificial seawater and 250 µl of 4% OsO_4. The solution was allowed to incubate for 10 min at room temperature, and then centrifuged at 3 000 g for 5 min. The pellet was washed twice by resuspending it in 1,5 ml of deionized water and subsequently centrifuged at 3 000 g for 10 min. The final pellet was resuspended in 1,5 ml of deionized water and the solution was pipetted onto a charged microscope slide (Huida, China). The Hybridization mix was prepared from one volume of probe working solution (50 µg/µl of 5′ Cy3.labeled Eub338 universal bacterial probe) and nine volumes of Hybridization buffer (Table 7). The probe solution was kept on ice in the dark throughout the whole process.

Tab. 7.: The composition of Hybridization buffer.

Stock reagent	Amount	Final concentration in the Hybridization buffer
5 M NaCl	360 µl	900 mM
1 M Tris/HCl pH 7,2	40 µl	20 mM
Formamide	700 µl	35%
10% SDS	2 µl	
Deionized H_2O	Add to 2 ml	

10 µl of Hybridization mix was added to each slide, subsequently covered by parafilm and placed into a darkened container. The slides were left to incubate at 46 ˚C for 90 min., then quickly rinsed by Washing buffer (Table 8), transferred into preheated Washing buffer, and incubated for 25 min at 48 ˚C. After incubation, each slide was rinsed by deionized water and was left to air-dry. Dry slides were mounted in Vecta Shield (Vector Laboratories, Inc.), covered with cover slides and stored in the dark at 4 ˚C until imaging.

Tab. 8.: The composition of Washing buffer.

Stock reagent	Amount	Final concentration in the Washing buffer
5 M NaCl	700 µl	0,07 M
1 M Tris/HCl	1 ml	20 mM
10% SDS	50 µl	0,01%
Deionized H$_2$0	Add to 50 ml	

3.5. Isolation of RNA

600 ml of grown *Hemistasia phaeocysticola* culture were harvested by centrifugation at 3 000 g for 10 min in 50 ml containers. Pellets were combined, 0,5 ml of TRIreagent was added and mixed by pipetting up and down several times. The solution was centrifuged at 12 000 g for 10 min at 4 ˚C. The supernatant was transferred into a new clean tube and kept at room temperature for 5 min. 100 µl of chloroform was added, the solution was thoroughly mixed by vigorous shaking and then left at room temperature for 15 min. The mixture was centrifuged at 12 000 g for 10 min at 4 ˚C. The top fraction was transferred into a new clean tube. Mixed with 250 µl of isopropanol and was left incubating at room temperature for 10 min. The solution was centrifuged at 12 000 g for 8 min at 4 ˚C and the supernatant was discarded. The RNA pellet was washed twice by adding 1 ml of 75% ethanol and subsequently centrifuged at 7 500 g for 5 min at 4 ˚C. The pellet was air-dried for approximately 10 min and was resuspended in TE buffer (10 mM Tris-HCl, 1 mM EDTA·Na$_2$, pH 8). The RNA yield was measured using Qubit and the integrity analyzed by BioAnalyzer. The above procedure was done for each of three samples per each of the three following category of culture – *H. phaeocysticola* in Hemi medium (Table 6), *H. phaeocysticola* in artificial seawater (36 g/l of sea salt [sera marin]) with no additional nutrients, and *H. phaeocysticola* in presence of live *Paracoccus* sp. with no additional nutrients.

3.6. Bioinformatics analysis

Total RNA from each of the above-mentioned categories in triplicates was sent for sequencing using Illumina eukaryotic messenger RNA-Seq., PE – 150 (pair-end reads 150 bp long) (Novogene). The quality of the raw RNA sequences was

confirmed via FastQC (Brown et al., 2017). We removed adapters and low-quality reads using Trimmomatic 0.39 (Bolger et al., 2014) with the following parameters: leading: 3, trailing: 3, slingwindow: 4:15, minlen: 36. FastQC was then used again as a control. The de novo transcriptome was generated using rnaSPAdes (Bushmanova et al., 2019) using default settings. To eliminate possible bacterial contamination, the transcripts were screened against nr NCBI database using the DIAMOND (Buchfink et al., 2014) and only those with at least one eukaryotic highest-scoring pair among the best three hits were considered for further analyses. Assembly was verified using Bowtie2 (Langmead & Salzberg, 2012), which aligned RNA-seq reads back to the transcriptome assembly to quantify read representation. More than 90% of the reads were mapped as correct pairs. We then used Salmon (Patro et al., 2017) to estimate transcript expression values and generate a transcript count matrix for genes and perform cross-sample normalization. The count matrices were used by EdgeR (McCarthy et al., 2012) to obtain differentially expressed genes, which were then compared between the above-mentioned three conditions. Next, analyze_diff_expr.pl script from Trinity package (Haas et al., 2013) was used to filter differentially expressed genes according to the following conditions: False Discovery Rate 0.001 and Fold Change c = 2, giving us a fourfold change. The obtained values differentially expressed genes passing above-mentioned conditions were used to generate Heatmaps. The expression values are plotted mean centered in log2 space. Negative values give us down-regulated genes and positive values give us up-regulated genes. Transcripts were annotated using OmicsBox (BioBam, Spain) through intersection of BLASTX against Swiss-Prot (Bairoch, 2000), InterProScan (Hunter et al., 2009), and EggNOG (Powell et al., 2011) databases. Annotated transcripts are grouped in Orthologus Groups of proteins (COG) categories. We used Fisher´s Exact Test (Enrichment analysis) to see only the most specific categories containing our annotated genes. (Khatri et al., 2002)

4. Results

4.1. Determination of bacterial prey for *Hemistasia phaeocysticola*

To determine the bacterial prey of *H. phaeocysticola* in our cultures, we amplified and sequenced the small bacterial ribosomal subunit (16S rRNA). The 16S rRNA amplicon (Figure 4) was sequenced by Eurofins Genomics. Resulting sequences were identified by nucleotide BLAST (Basic Local Alignment Search Tool) through the NCBI GenBank web interface. The highest scoring pairs were undescribed species from the genus *Paracoccus*. The greatest percentage of identity was 98,56% to *Paracoccus marcusii* and more than 90% to other species in the genus of *Paracoccus*.
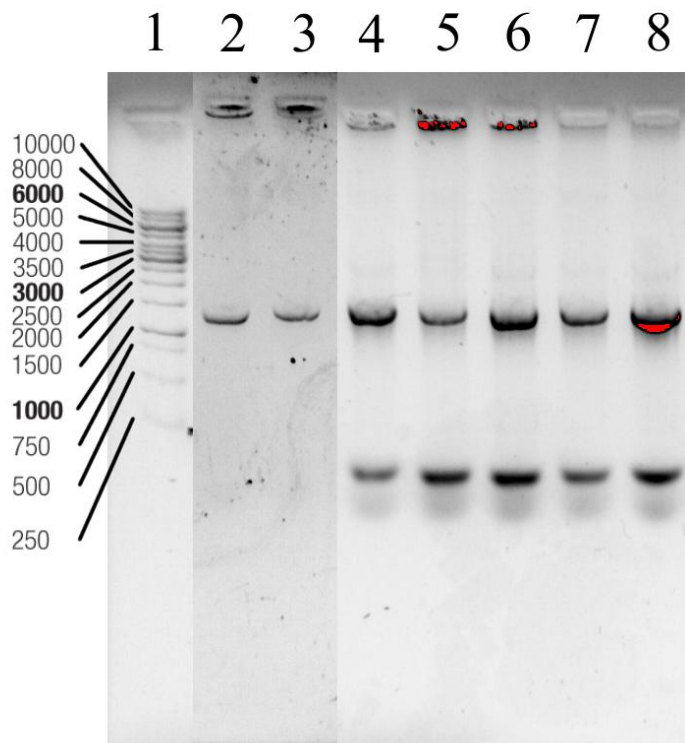

Fig. 4.: Agarose gel electrophoresis of all the steps for uncovering the genus of bacteria used as a prey for *Hemistasia phaeocysticola*. 1 is 1kb ladder; 2-3 are products of PCR from 2 separate grown colonies of unknown bacteria plated on Marine broth plate; 4-8 are PCR products of transformed E. Coli with 16s rRNA gene.

4.2. Growth of *Hemistasia phaeocysticola* preying on live *Paracoccus* sp.

Figures 5, 6 show light microscopy using Differential Interference Contrast (DIC) of *Hemistasia phaeocysticola* cells preying on live *Paracoccus* sp. bacteria. A small portion of *H. phaeocysticola* cells had difficulty dividing (Figure 5, 6) while fed with live bacterial cells. Each time this was observed, unusually large, full vacuoles were shared. A similar phenotype was observed in other diplonemids, namely *Lacrimia 1601* and *Lacrimia 1608*, in which cases large full vacuoles were observed along with difficulty dividing while in medium during the logarithmic phase of growth (Tashyreva, pers.comm.). As far as I know, this phenotype was not observed in other diplonemids without a single large vacuole. What exactly caused this problem is currently unknown.


Fig. 5.: DIC light microscopy of *Hemistasia phaeocysticola* preying on live *Paracoccus* sp. bacteria. The top cell represents well fed cell with big vacuole full of bacteria. The bottom cell represents cell in the process of dividing, however is unable to divide properly due to overly large and full vacuole resulting in two-cell live hybrid.
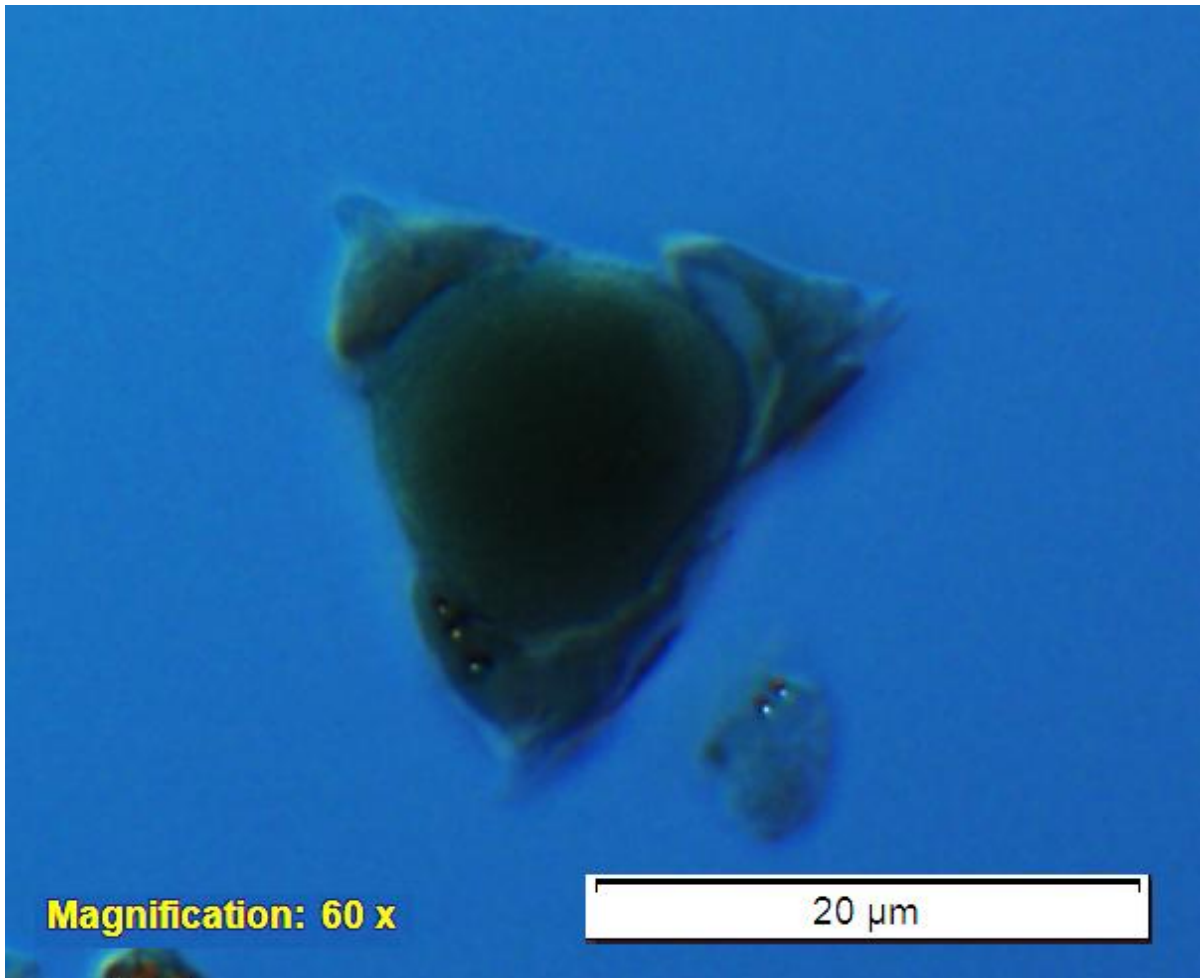
Fig. 6.: DIC light microscopy of *Hemistasia phaeocysticola* preying on live *Paracoccus* sp. bacteria. It is unable to properly divide due to overly large and full vacuole resulting in free-cell live hybrid.

4.3. Growth curve of *Hemistasia phaeocysticola* preying on live *Paracoccus* sp.

To determine whether *Hemistasia phaeocysticola* can survive and thrive on *Paracoccus* sp. as the only source of nutrients a growth curve (Figure 7, 8) was prepared. In the graph (Figure 7) growth curves from three cultures prepared and grown under identical conditions and one negative growth control are shown. The first three to four days are an adjustment period during which *H. phaeocysticola* adjusted to a new source of nutrients. They are followed by a faster growth rate during the next few days. After that, *H. phaeocysticola* cells started to die, possibly from the lack of available bacteria. The next graph (Figure 8), are the exact same growth curves with the addition of positive control. The graphs were split to properly show the differences in the growth rate in medium and on bacteria in regards to negative control. The growth rate of *H. phaeocysticola* preying on *Paracoccus* sp. is nowhere near as high as the growth rate of *H. phaeocysticola* in Hemi medium (Table 6).
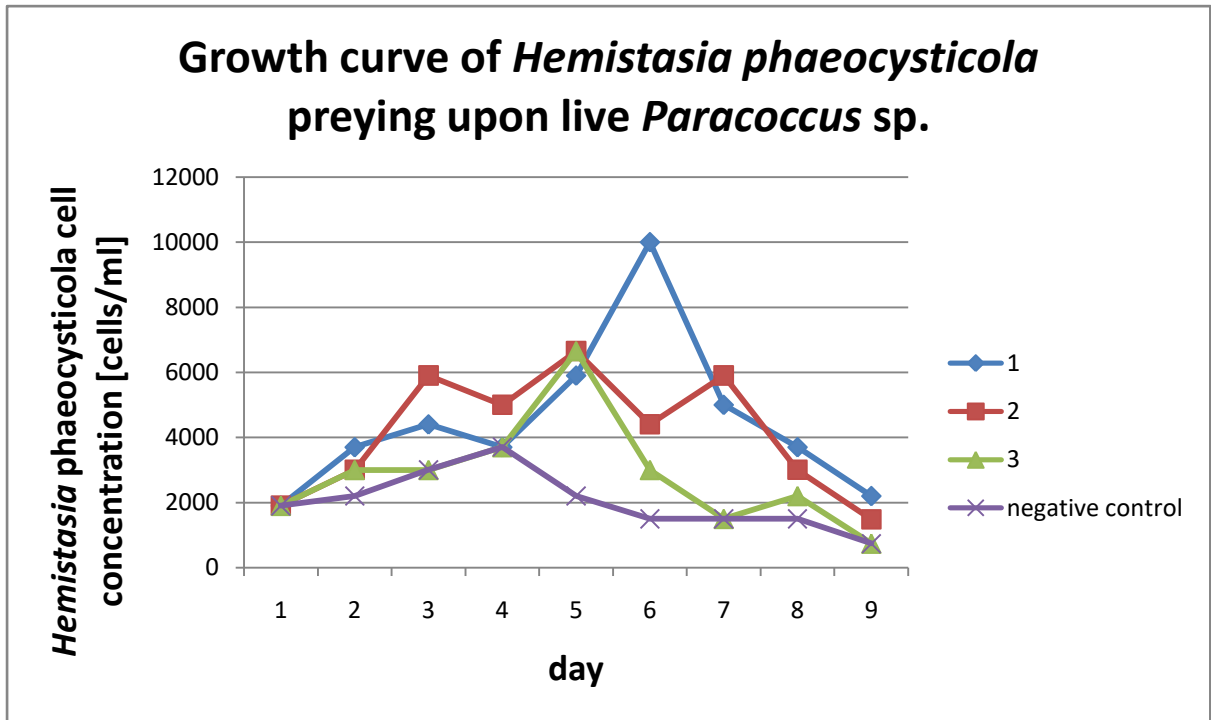
Fig. 7.: The growth curve of *Hemistasia phaeocysticola* preying on live *Paracoccus* sp. bacteria. The graphs 1,2,3 represent 3 distinct growth curves cultures of *H. phaeocysticola* preying on live *Paracoccus* sp. and the negative control graph represent growth curve of *H. phaeocysticola* in artificial seawater. The samples were taken and measured every 24 hours.
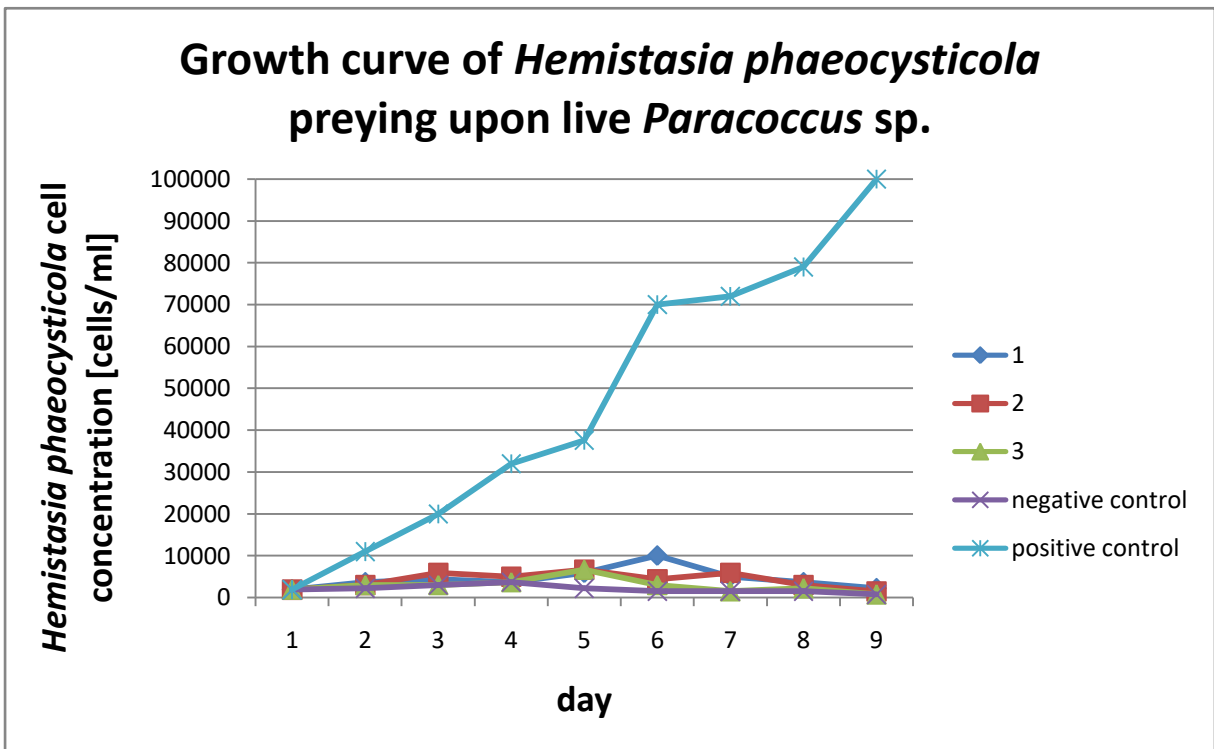
Fig. 8.: The growth curve of *Hemistasia phaeocysticola* preying on live *Paracoccus* sp. bacteria. The graphs 1,2,3 represent 3 distinct growth curves cultures of *H. phaeocysticola* preying on live *Paracoccus* sp. The negative control graph represent growth curve of *H. phaeocysticola* in artificial seawater. The positive control graph represent growth curve of *H. phaeocysticola* in Hemi medium. The samples were taken and measured every 24 hours.

4.4. Fluorescence microscopy of *Paracoccus* sp. inside vacuoles of *Hemistasia phaeocysticola*

Another source of evidence of bacteriovory in *Hemistasia phaeocysticola* is fluorescence microscopy. Using the suitable probe, it enables visualizing bacteria cells inside the vacuole of *H. phaeocysticola*. The figure 9 shows oval to circular shaped *Paracoccus* sp. bacteria cells sticking together forming clumps and chains visualized by phase contrast in grey; universal bacterial probe in red; DAPI, which binds to any DNA in blue; composite of all three previous figures together to show co-localization.
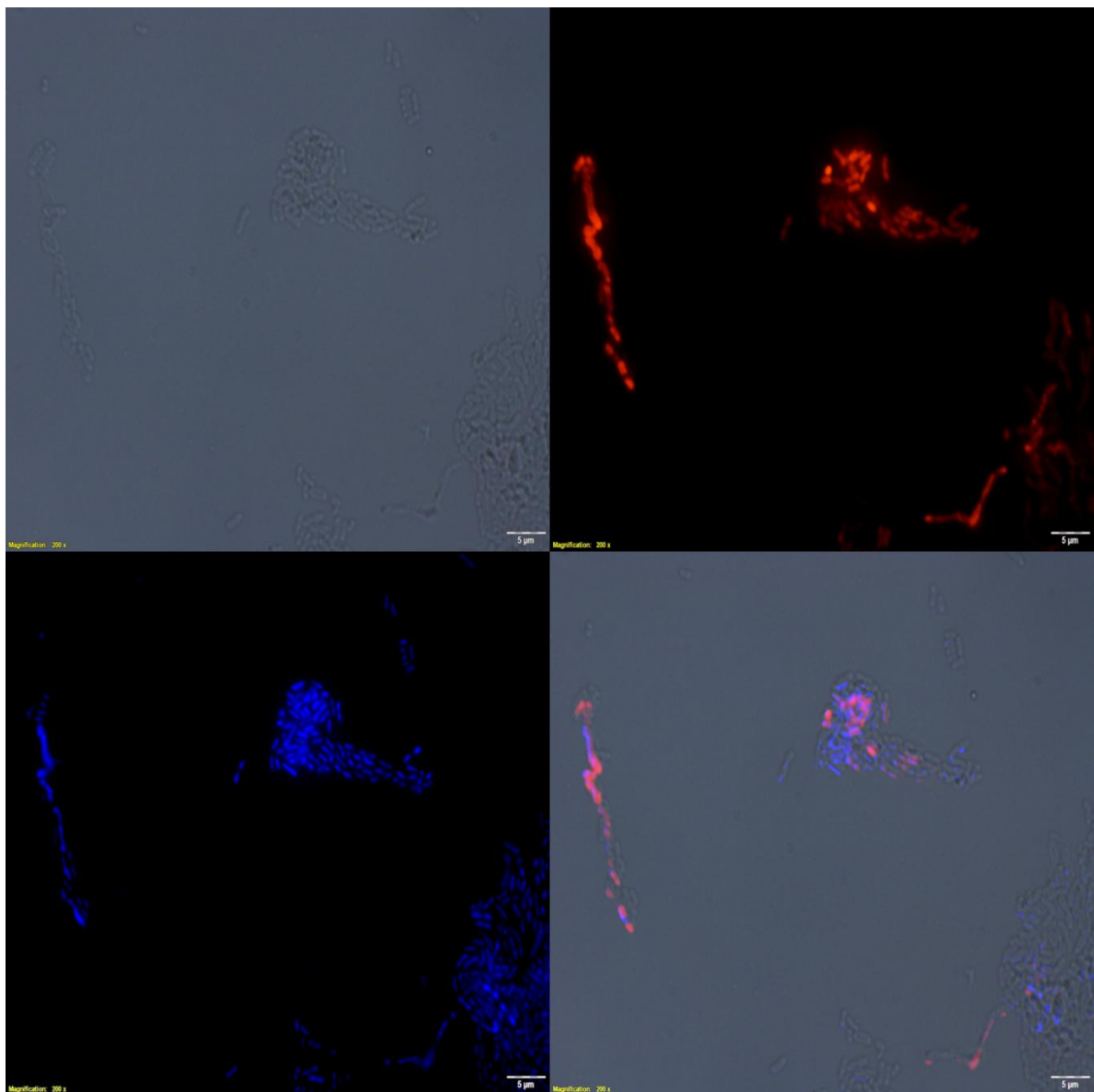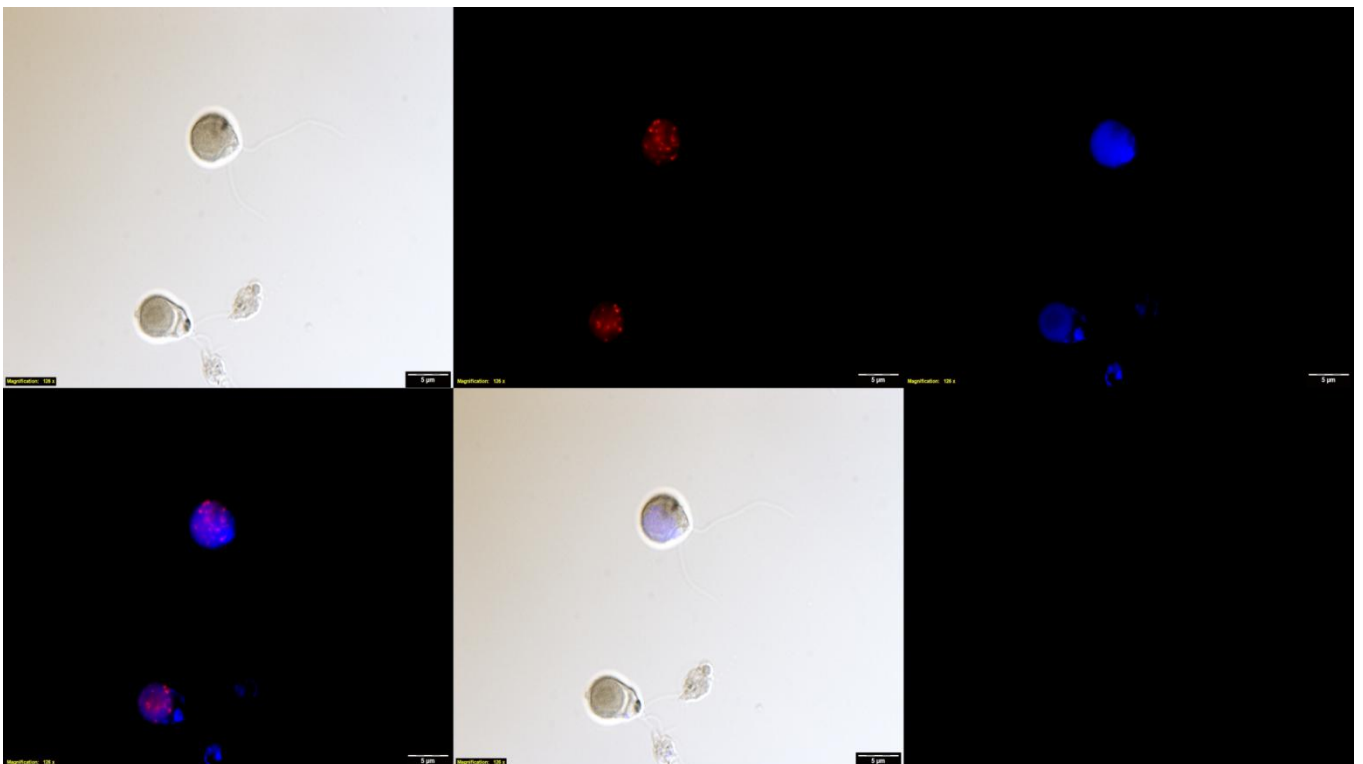

Fig. 9.: The montage of figures depicting from top left to bottom right: phase contrast in grey of Pracoccus sp.; universal bacterial probe in red; DAPI in blue; composite of all previous figures combined for co-localization.

Once we were assured that the bacterial probe is working as intended, we visualized *H. phaeocysticola* under the fluorescence microscope using the same wavelength as was used for the bacterial probe to rule out possible autofluorescence. Figure 10 shows phase contrast in grey; slight autofluorescence can be seen in red coming from the vacuole, however faint enough to not cause any false results. DAPI in blue binds to any DNA and shows genomic and mitochondrial DNA of *H. phaeocysticola*; composite of all previous figures shows co-localization.
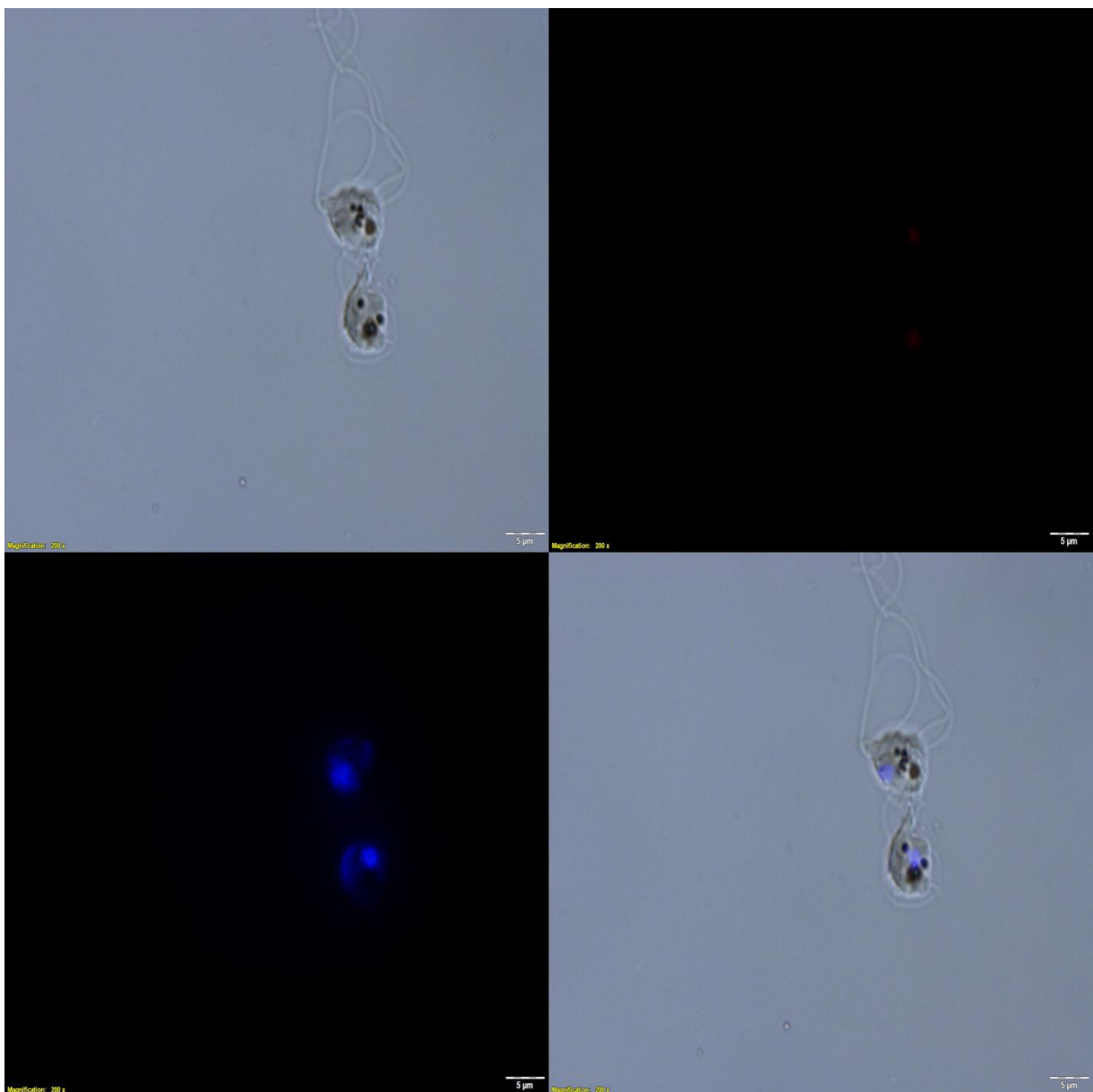
Fig. 10.: The montage of figures depicting from top left to bottom right: phase contrast in grey of *Hemistasia phaeocysticola*; autofluorescence in red of *H. phaeocysticola* under the same wavelength as used for visualizing universal bacterial probe used above; DAPI in blue binding to the DNA of *H. phaeocysticola*; composite of all previous figures to visualize co-localization.

In the next step, we visualized *H. phaeocysticola* cells preying on live *Paracoccus* sp. under the fluorescence microscope. Figure 11 shows phase contrast in grey; universal bacterial probe in red indicates the shapes of still intact and partially digested bacteria inside vacuole. DAPI binding to all DNA visualizing the genomic and mitochondrial DNA of *H. phaeocysticola* as well as to the DNA of still intact and partially digested *Paracoccus* sp. bacteria inside the vacuole of *H. phaeocysticola.* Composite figure shows co-localization of the previous figures which points out that bacteria are inside the vacuole and are in the process of digesting.

Fig. 11.: The montage of figures depicting from top left to bottom right: phase contrast in grey of *Hemistasia phaeocysticola*; universal bacterial probe in red of; DAPI in blue binding to the genome and mitochondrial DNA of *H. phaeocysticola* and to the DNA of still intact and partially digested *Paracoccus* sp. bacteria inside the vacuole of *H. phaeocysticola*; the composite of two previous figures – universal bacterial probe and DAPI to show co-localization of bacteria cells with bacterial DNA inside the vacuole of *H. phaeocysticola*; the composite of all previous figures showing co-localization.

4.5. Bioinformatics analysis

We obtained the following number of raw reads from Illumina eukaryotic messenger RNA-Seq (Novogene): *Hemistasia phaeocysticola* in Hemi medium (Table 6) each of the triplicate (forward + reverse): 153 993 626, 119 040 960, 137 382 360; *H. phaeocysticola* in seawater medium without nutrients each of the triplicate (forward + reverse): 114 886 800, 88 747 836, 85 228 384; *H. phaeocysticola* preying on *Paracoccus* sp. each of the triplicate (forward + reverse): 116 200 084, 100 115 786, 94 810 258. RnaSPAdes assembled a total of 135 416 transcripts from raw sequences with median contig length of 1 190 and contig N50 (length of the shortest contig at 50% of the transcriptome length) of 2 923. From the total number of assembled transcripts, we filtered out all putative non-eukaryotic transcripts and obtained 43 094 transcripts which were used in following analyses. The median coverage of eukaryotic transcripts was 405,6. From all eukaryotic transcripts 37 392 were annotated leaving 5 702 not annotated.

RNA-seq data of differentially expressed genes obtained with EdgeR were visualized by volcano plots (Figure 12, 13, 14) representing fold change versus false discovery rate (FDR value). Each triplicate set of sequences was compared to each other. From bottom to top of the plot we see statistically less to more highly significant genes and on the logFC axis from zero to the left of the plot we see down-regulated genes and from zero to the right of the plot we see up-regulated genes.

Hemi medium versus Seawater without nutrients
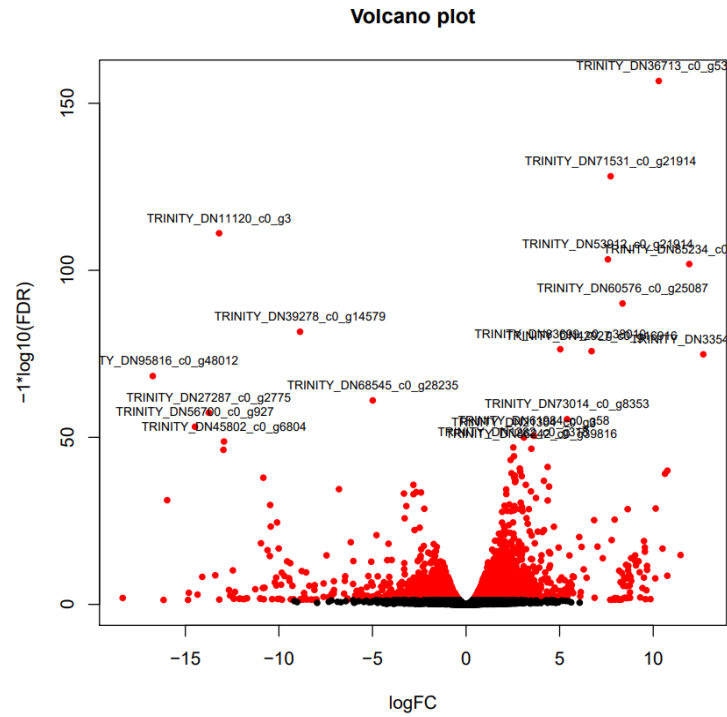
Fig. 12.: Volcano plot showing fold change versus statistical significance (FDR value) of *Hemistasia phaeocysticola* in Hemi medium (Table X) compared to *H. phaeocysticola* maintained in artificial seawater (sera marin) without nutrients for 48 hours. From bottom to top of the graph we see less to higher statistically significant genes and from left to right of the graph we see down-regulated to up-regulated genes.

Bacterial prey vs Seawater medium
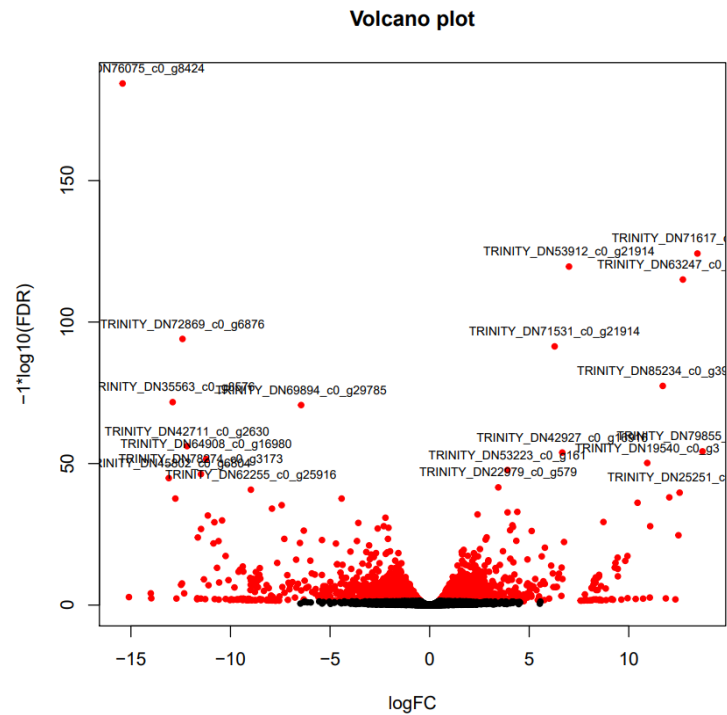
Fig. 13.: Volcano plot showing fold change versus statistical significance (FDR value) of *Hemistasia phaeocysticola* preying on *Paracoccus* sp. bacteria compared to *H. phaeocysticola* kept in artificial seawater (sera marin) without nutrients for 48 hours. From bottom to top of the graph we see less to higher statistically significant genes and from left to right of the graph we see down-regulated to up-regulated genes.

Hemi medium vs Bacterial prey
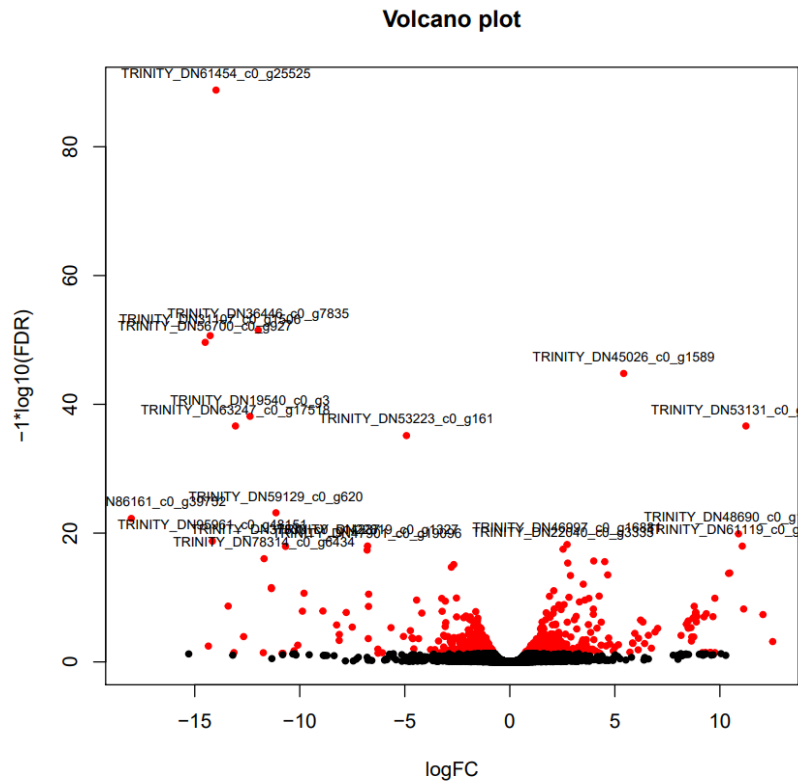
Volcano plot


Fig. 14.: Volcano plot showing fold change versus statistical significance (FDR value) of *Hemistasia phaeocysticola* in Hemi medium (Table X) compared to *H. phaeocysticola* preying on *Paracoccus* sp. bacteria. From bottom to top of the graph we see less to higher statistically significant genes and from left to right of the graph we see down-regulated to up-regulated genes.

Next, from the differentially expressed genes obtained using EdgeR, we generated Heatmap (Figure 15), showing down-regulated genes in shades of violet and up-regulated genes in shades of yellow. H481, H482, H483 represent three distinct cultures of *Hemistasia phaeocysticola* maintained in environment without nutrients for 48 hours. HB1, HB2, HB3 represent three distinct cultures of *H. phaeocysticola* preying on *Paracoccus* sp. bacteria. HPM1, HPM2, HPM3 represent three distinct cultures of *H. phaeocysticola* maintained in Hemi medium (Table 6). As seen from the Heatmap, samples cluster according respective sampling condition, suggesting that the distribution of differentially expressed genes is sufficiently similar between triplicates. Also, samples grown on bacteria and nutrition-rich medium are more similar to each other, than to samples grown without presence of nutritions.
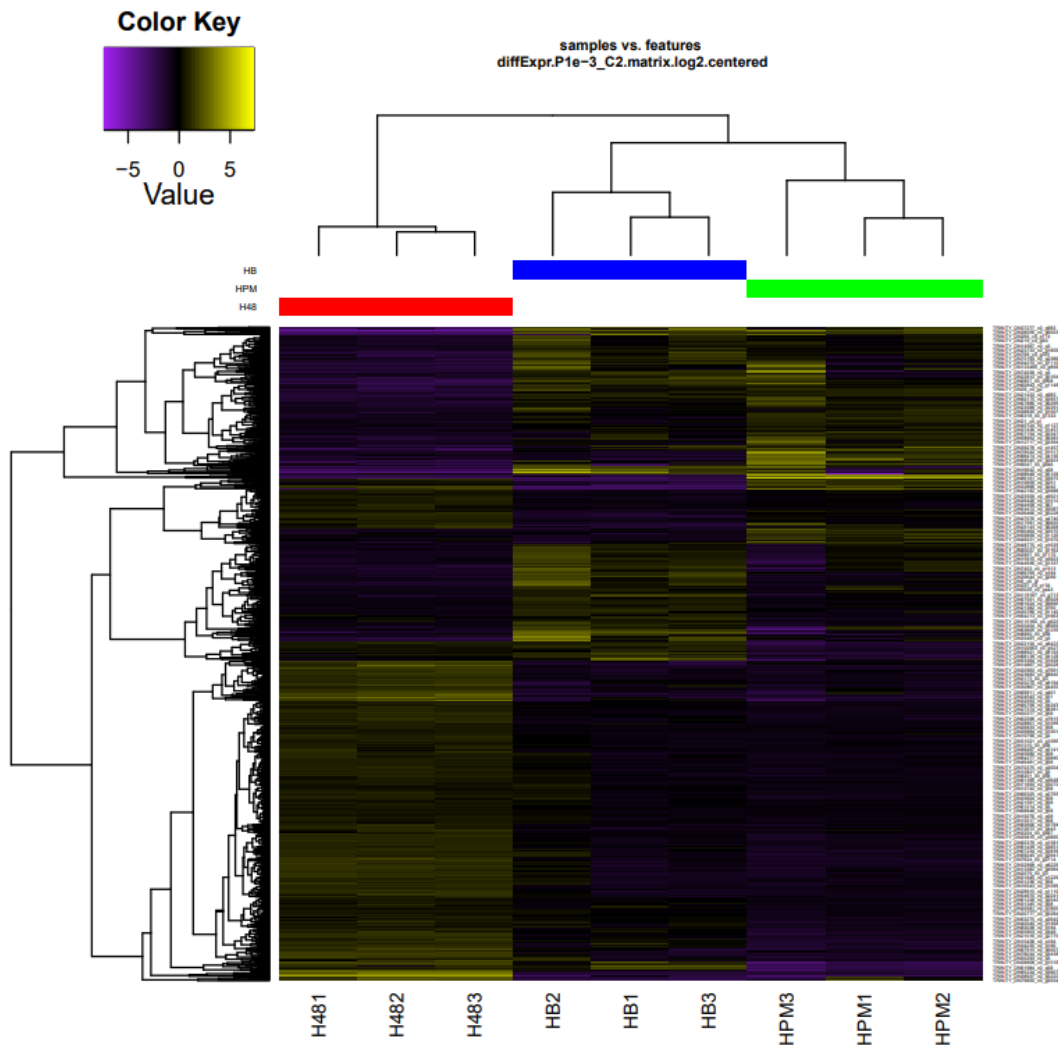
Fig. 15.: Heatmap generated using DESeq2 shows down-regulated genes in shades of yellow and up-regulated genes in shades of violet. H481, H482, H483 represent three distinct cultures of *Hemistasia phaeocysticola* kept environment without nutrients for 48 hours. HB1, HB2, HB3 represent three distinct cultures of *H. phaeocysticola* preying on *Paracoccus* sp. bacteria. HPM1, HPM2, HPM3 represent three distinct cultures of *H. phaeocysticola* maintained in Hemi medium (Figure X).

Using OmicsBox we created spreadsheets of annotated differentially expressed genes which can be found as Supplement 1. Each triplicate set of sequences was compared to the others. The most relevant sets of annotated differentially expressed genes were sorted in Tables 9, 10, 11. Table 9 shows the comparison of the annotated differentially expressed genes of *H. phaeocysticola* preying on *Paracoccus* sp. bacteria versus *H. phaeocysticola* maintained in the Hemi medium (Table 6). Noteworthy is that the number of up-regulated and down-regulated genes in all Orthologous Groups of proteins (COG) categories in this comparison is significantly smaller than in the other two comparisons comparing *H. phaeocysticola* with a nutrient source to *H. phaeocysticola* maintained without nutrients for 48 hours.

Tab. 9.: The comparison of annotated differentially expressed genes of *Hemistasia phaeocysticola* preying on *Paracoccus* sp. bacteria versus *H. phaeocysticola* in Hemi medium (Table 6). The General information section shows the basic information for down-regulated and up-regulated annotated differentially expressed genes. The Database of Clusters of Orthologous Groups of proteins (COG) Categories Distribution section shows the number and percentage of down-regulated and up-regulated annotated differentially expressed genes belonging to basic categories.

HB vs HPM					
General information			**COG Categories Distribution**		
	down-regulated	**up-regulated**		**down-regulated**	**up-regulated**
Total amount of input sequences:	86	63	**Information Storage and Processing:**	20 / 26.67%	8 / 15.09%
Average length:	2926.0	2005.0	**Cellular Processes and Signalling:**	21 / 28.0%	22 / 41.51%
Number of GO annotated sequences:	35	31	**Metabolism:**	19 / 25.33%	17 / 32.08%
Number of GO annotations:	340	461	**Function unknown**	12 / 16.0%	6 / 11.32%
Average GOs per sequence:	9.71	14.87			

Tab. 10.: The comparison of annotated differentially expressed genes of *Hemistasia phaeocysticola* preying on *Paracoccus* sp. bacteria versus *H. phaeocysticola* maintained in artificial seawater (sera marin) for 48 hours. The General information section shows the basic information for down-regulated and up-regulated annotated differentially expressed genes. The Database of Clusters of Orthologous Groups of proteins (COG) Categories Distribution section shows the number and percentage of down-regulated and up-regulated annotated differentially expressed genes belonging to basic categories.

HB vs H48					
General information			**COG Categories Distribution**		
	down-regulated	**up-regulated**		**down-regulated**	**up-regulated**
Total amount of input sequences:	240	310	**Information Storage and Processing:**	36 / 18.75%	35 / 12.64%
Average length:	2646.0	5355.0	**Cellular Processes and Signalling:**	65 / 33.85%	116 / 41.88%
Number of GO annotated sequences:	100	143	**Metabolism:**	53 / 27.6%	65 / 23.47%
Number of GO annotations:	1742	2063	**Function unknown**	25 / 13.02%	44 / 15.88%
Average GOs per sequence:	17.42	14.43			

Tab. 11.: The comparison of annotated differentially expressed genes of *Hemistasia phaeocysticola* in Hemi medium (Table 6) versus *H. phaeocysticola* maintained in artificial seawater (sera marin) for 48 hours. The General information section shows the basic information for down-regulated and up-regulated annotated differentially expressed genes. The Database of Clusters of Orthologous Groups of proteins (COG) Categories Distribution section shows the number and percentage of down-regulated and up-regulated annotated differentially expressed genes belonging to basic categories.

HPM vs H48					
General information			**COG Categories Distribution**		
	down-regulated	**up-regulated**		**down-regulated**	**up-regulated**
Total amount of input sequences:	536	209	**Information Storage and Processing:**	112 / 23.48%	29 / 15.51%
Average length:	2547.0	3569.0	**Cellular Processes and Signalling:**	164 / 34.38%	78 / 41.71%
Number of GO annotated sequences:	193	111	**Metabolism:**	93 / 19.5%	48 / 25.67%
Number of GO annotations:	3183	1513	**Function unknown**	79 / 16.56%	17 / 9.09%
Average GOs per sequence:	16.49	13.63			

To step up from the individual genes, we performed Fisher´s Exact enrichment test, which compares the whole pathways or biological processes, and help us to draw out meaningful picture from the gene-based analyses. The results are shown as Word Cloud diagrams. Figure 16, 17 shows the Word Cloud diagram of annotated differentially expressed genes of *H. phaeocysticola* comparing starving cells maintained in artificial seawater (sera marin) without nutrients for 48 hours to the annotated differentially expressed transcripts of fed *H. phaeocysticola* maintained in Hemi medium (Table 6) and preying on *Paracoccus* sp. bacteria. The text represents the name of COG group containing annotated differentially expressed genes and the size of the font shows the number and significance of genes annotated to that particular group.

The figure 16 shows Word Cloud diagram of most significant up-regulated transcripts of *H. phaeocysticola* maintained in artificial seawater (sera marin) without nutrients for 48 hours compared to fed *H. phaeocysticola* maintained in Hemi medium (Table 6) and preying on bacteria. It shows scavenger receptor activity group of proteins is highly

up-regulated. Into this category belong proteins related to endocytosis, phagocytosis, adhesion, and signalling leading to the degradation of substances (PrabhuDas et al., 2017). This suggests that the starving cells actively transcribe proteins able to recognize, eliminate, and or digest potentially harmful or nutritionally useful substances.

Fig. 16.: The Word Cloud diagram shows the most significant COG groups containing annotated up-regulated transcripts of *Hemistasia phaeocysticola* maintained in artificial seawater (sera marin) without nutrients for 48 hours compared to fed *H. phaeocysticola* maintained in Hemi medium (Table 6) and preying on *Paracoccus* sp. bacteria.

The Figure 17 shows Word Cloud diagram of most significant up-regulated transcripts of fed *H. phaeocysticola* maintained in Hemi medium (Table 6) and preying on bacteria compared to *H. phaeocysticola* maintained in artificial seawater (sera marin) without nutrients for 48 hours. The up-regulated transcripts suggest that fed *H. phaeocysticola* invest more into maintaining the cell-homeostasis, structural proteins, and proteins helping modify other proteins like protein kinases.
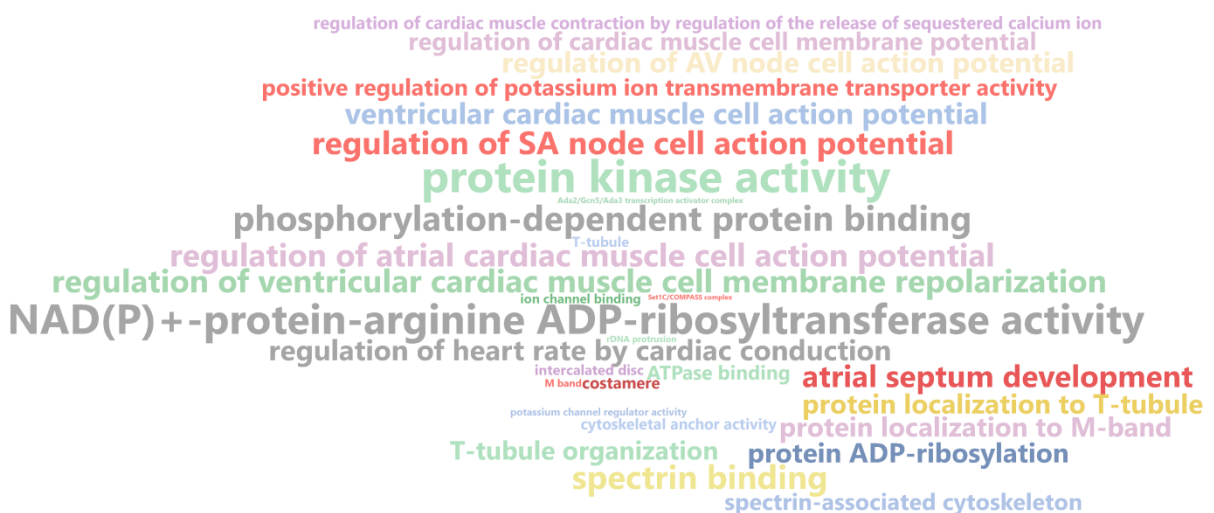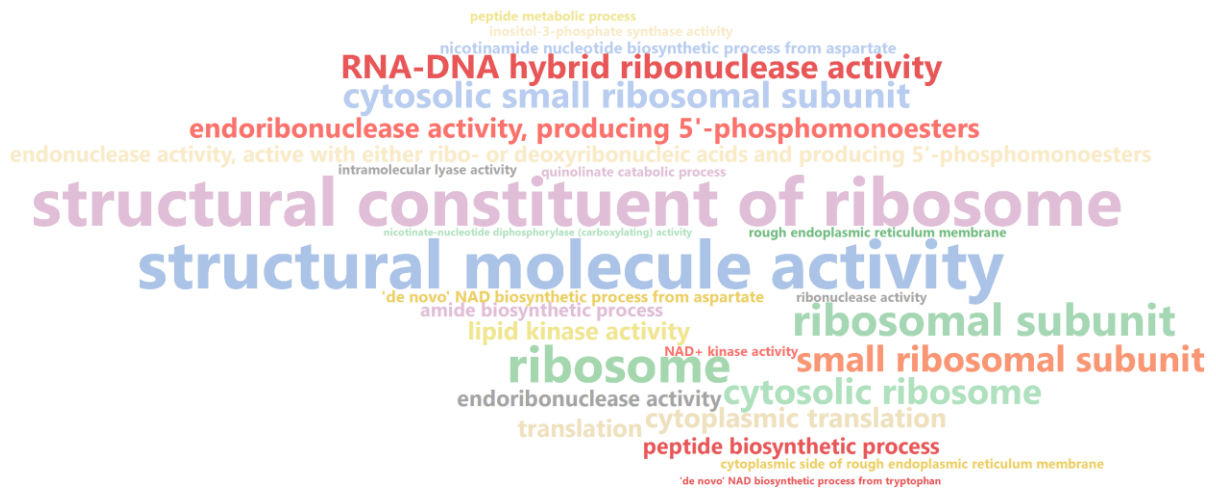

Fig. 17.: The Word Cloud diagram shows the most significant COG groups containing annotated up-regulated transcripts of fed *Hemistasia phaeocysticola* maintained in Hemi medium (Table 6) and preying on *Paracoccus* sp. bacteria compared to *H. phaeocysticola* maintained in artificial seawater (sera marin) without nutrients for 48 hours.

31

The Figure 18 shows Word Cloud diagram of most significant up-regulated transcripts of *H. phaeocysticola* preying on *Paracoccus* sp. bacteria compared to *H. phaeocysticola* maintained in Hemi medium (Table 6). The up-regulated transcripts suggest that *H. phaeocysticola* preying on bacteria invest more into structural proteins, ribosomes, and endoribonuclease proteins.
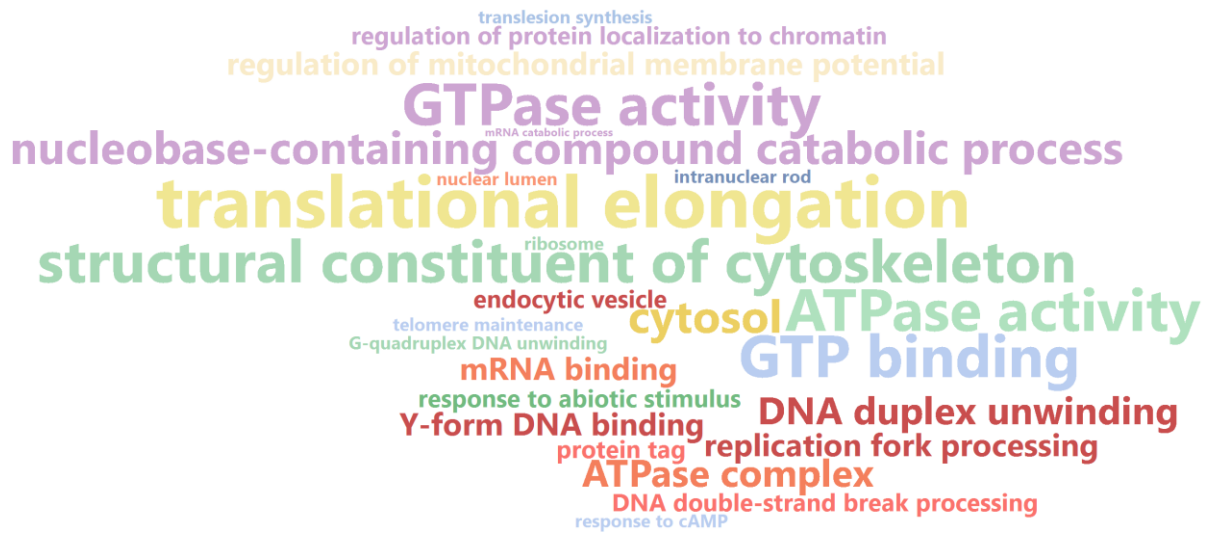

Fig. 18.: The Word Cloud diagram shows the most significant COG groups containing annotated up-regulated transcripts of *Hemistasia phaeocysticola* preying on preying on *Paracoccus* sp. bacteria compared to *H. phaeocysticola* maintained in Hemi medium (Table X).

The Figure 19 shows Word Cloud diagram of most significant up-regulated transcripts of *H. phaeocysticola* maintained in Hemi medium (Table 6) compared to *H. phaeocysticola* preying on *Paracoccus* sp. bacteria. *H. phaeocysticola* maintained in nutrition-rich Hemi medium (Table 6) invest more into maintaining the cell, structure, translation and the apparatus to directly bind nutrients and transform them easily into ATP and GTP.

Fig. 19.: The Word Cloud diagram shows the most significant COG groups containing annotated up-regulated transcripts of *H. phaeocysticola* maintained in Hemi medium (Table X) compared to *Hemistasia phaeocysticola* preying on preying on *Paracoccus* sp. bacteria.

5. Discussion

Heterotrophic flagellates are recognized as main bacterial grazers throughout ocean ecosystems, although not all heterotrophic flagellates are bacteriovores. (Lukeš et al., 2015) There is significant variability in prey selection and intake of nutrients. Genus and even species-specific feeding habits are commonly found. The feeding habits of increasingly narrower groups of protists are being uncovered to properly understand the intricate food webs in marine ecosystems. (Boenigk & Arndt, 2002) Currently, there are two confirmed bacteriovores among diplonemids, namely *Diplonema japonicum* and *Rhynchopus humris*. (Prokopchuk, pers.comm.)

To detect bacteriovory in *Hemistasia phaeocysticola*, we used several strategies. First, we constructed a growth curve (Figure 7, 8) which showed us that compared to the negative control consisting of *H. phaeocysticola* in artificial seawater without nutrients, there is significant growth of the culture, but compared to the positive control consisting of *H. phaeocysticola* in the Hemi medium (Table 6), the growth rate of *H. phaeocysticola* on live *Paracoccus* sp. bacteria is lower. This suggests that live *Paracoccus* sp. bacteria are possible as a nutrient source, however not optimal as the sole source. Moreover, as shown in DIC light microscopy (Figure 5, 6), *Paracoccus* sp. as the sole source of nutrients can lead to abnormalities in reproduction. *Paracoccus marcusii* and other *Paracoccus* species are known to produce considerable quantities of carotenoids including astaxanthin, giving them characteristic orange colour (Harker et al., 1998). Fluorescence microscopy shows the bacterial cells in the vacuole during the digestion process, which together with the previous point suggests that *H. phaeocysticola* can indeed prey on *Paracoccus* sp. and probably on other bacterial species. In my opinion it would be beneficial to repeat the above-mentioned experiments with dead bacterial cells and additionally measure bacterial concentration. Previous observations of nutrient sources consisting of diatoms, haptophytes, dinoflagellates and even dead *H. phaeocysticola* cells (Elbrächter et al., 1996), along with our observation of bacteriovory suggest to us that *H. phaeocysticola* might be a generalist with a wide range of nutrient sources, however, more recent observations and experiments are needed to answer the question of what exactly is and is not a nutrient source for *H. phaeocysticola*.

Preliminary, unpublished experiments showed a much lower frequency of abnormalities during the reproduction of *H. phaeocysticola* while preying on dead *Paracoccus* sp. Other unpublished results (Prokopchuk, pers.comm.) showed an inability to grow while preying on a mixture of dead bacteria, suggesting that *H. phaeocysticola* can only thrive when fed on only certain bacteria or only under certain conditions. It would be interesting to find out to what extent *H. phaeocysticola* is a bacteriovore.

The number of differentially expressed genes may likely be higher as we removed some of them to avoid potential bacterial contamination. Also, due to lack of annotated genome of *H. phaeocysticola*, it is impossible to differentiate genes between possible bacterial contamination and horizontal gene transfer. Although *H. phaeocysticola* has been described for a relatively long time, no transcriptomic analyses have yet been published. Other diplonemids have also only recently begun to be studied at the genetic level, so we consider the set of annotated genes obtained to be a good start for further bioinformatic studies on diplonemids and *H. phaeocysticola* in particular. Differentially expressed gene analyses showed similarities in up- and down-regulated genes (Figure 15) when *H. phaeocysticola* was able to obtain nutrients. The differences were much greater when differentially expressed genes from cultures of *H. phaeocysticola* that were able to gain nutrients were compared with those from starved cultures. Of all the COG groups that contained annotated differentially expressed transcripts (Table 9, 10, 11), the largest differences across all food source comparisons were in Metabolism, and Cellular Processes and Signalling categories containing most of the annotated differentially expressed genes. Changes in genes involved in metabolism are to be expected. The cell only needs certain genes to be active to process available nutrients, which can vary greatly depending on the source. Cellular Processes and Signalling category might contain genes that control phenotypic changes in shape and behaviour, depending on whether a cell can only take in nutrients from the environment or whether a cell needs to actively hunt prey. Tracking down these genes could help us understand the behaviour of *H. phaeocysticola*.

Comparing the up-regulated genes of *H. phaeocysticola* maintained in artificial seawater (sera marin) for 48 hours without available nutrients to fed *H. phaeocysticola* in Hemi medium (Table 6) and preying on *Paracoccus* sp. bacteria, the most prominent group of genes was Scavenger Receptor Activity as shown in Figure 16. To this category belong genes responsible for delivering large variety of ligands to the cell via endocytosis. This is to be expected as these genes may help in the uptake of nutrients from the medium

or bacteria. Viewed from the other side of the comparison, the most significant up-regulated transcripts of fed *H. phaeocysticola* maintained in Hemi medium (Table 6) and preying on bacteria compared to *H. phaeocysticola* maintained in artificial seawater (sera marin) without nutrients for 48 hours. The up-regulated groups' (Figure 17) show us that fed *H. phaeocysticola* invest more into proteins helping with development of the cell, regulatory, and modifying categories like protein kinases. This leads us to believe that while *H. phaeocysticola* is in an environment rich in nutrients, it focuses more on developing the cell structures. It further implies that the phenotype of *H. phaeocysticola* is significantly altered by many regulatory genes in an environment with and without nutrients.

There are not many published research articles examining free-living flagellated heterotrophic protists using bioinformatic approaches similar to those described in this thesis, which is further complicated by the lack of a widely accepted suitable model organism. One notable article is research conducted on proven bacteriovore, the heterotrophic flagellate *Cafeteria burkhardae* (Massana et al., 2020). The study consisted of a transcriptomic study of *C. burkhardae* grown on the flavobacterium *Dokdonia* sp. They compared gene expression between exponential and stationary phases supplemented with three starvation and dilution phases. Their studies revealed 2056 differentially expressed genes between exponential and stationary phases, which were annotated in a number of prominent categories, including some highly expressed phagocytosis genes (peptidases, proton pumps) that could potentially be used to target these processes in marine ecosystems. (Massana et al., 2020) Currently, there is a gap in sufficient research on the molecular mechanisms of predation in heterotrophic flagellates that could allow us to better understand the coexistence of protists in marine ecosystems.

6. Summary

In this thesis we measured the ability of *Hemistasia phaeocysticola* to grow on bacteria using growth curve in distinct growth conditions (medium, seawater medium, bacteria) and fluorescence microscopy. Next, we de novo assembled transcriptome and analyzed differentially expressed genes from all three growth conditions.

H. phaeocysticola exhibited a significant ability to grow on *Paracoccus* sp. bacteria, although to much lesser degree than in medium. We therefore conclude *Paracoccus* sp. is not an optimal food source for *H. phaeocysticola*, which is probably not specialized bacteriovore. Transcriptomic profiles correlate more closely between samples with an available source of nutrients than to cultures without nutrients. Fisher´s Exact enrichment test of *H. phaeocysticola* maintained in seawater without nutrients up-regulated transcripts of scavenger receptor activity proteins which suggest up-regulation of proteins able to recognize, eliminate, or digest potentially harmful or nutritionally useful substances. In comparison up-regulated transcripts of fed *H. phaeocysticola* suggest that fed cells invest more into proteins helping with development of the cell and proteins helping modify other proteins like protein kinases than starving cells. Up-regulated annotated transcripts of *H. phaeocysticola* preying on *Paracoccus* sp. bacteria suggest that cells preying on bacteria invest more into structural, ribosomal, and endoribonuclease proteins in comparison to cells maintained rich medium. In comparison up-regulated transcripts of *H. phaeocysticola* maintained in rich medium suggest that those cells invest more into structural proteins, apparatus to directly bind easily digested nutrients and easily transform them into ATP and GTP.

7. References

- Azam, F., & Malfatti, F. (2007). Microbial structuring of marine ecosystems. *Nature Reviews Microbiology*, *5*(10), 782–791. https://doi.org/10.1038/nrmicro1747

- Bairoch, A. (2000). The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Research*, *28*(1), 45–48. https://doi.org/10.1093/nar/28.1.45

- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

- Boenigk, J., & Arndt, H. (2002). Bacterivory by heterotrophic flagellates: community structure and feeding strategies. *Antonie van Leeuwenhoek*, *81*(1/4), 465–480. https://doi.org/10.1023/a:1020509305868

- Brown, J., Pirrung, M., & McCue, L. A. (2017). FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. *Bioinformatics*, *33*(19), 3137–3139. https://doi.org/10.1093/bioinformatics/btx373

- Buchfink, B., Xie, C., & Huson, D. H. (2014). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, *12*(1), 59–60. https://doi.org/10.1038/nmeth.3176

- Bushmanova, E., Antipov, D., Lapidus, A., & Prjibelski, A. D. (2019). rnaSPAdes: a de novo transcriptome assembler and its application to RNA-Seq data. *GigaScience*, *8*(9). https://doi.org/10.1093/gigascience/giz100

- Calbet, A., & Landry, M. R. (2004b). Phytoplankton growth, microzooplankton grazing, and carbon cycling in marine systems. *Limnology and Oceanography*, *49*(1), 51–57. https://doi.org/10.4319/lo.2004.49.1.0051

- de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahe, F., Logares, R., Lara, E., Berney, C., Le Bescot, N., Probert, I., Carmichael, M., Poulain, J., Romac, S., Colin, S., Aury, J. M., Bittner, L., Chaffron, S., Dunthorn, M., Engelen, S., . . . Velayoudon, D. (2015). Eukaryotic plankton diversity in the sunlit ocean. *Science*, *348*(6237), 1261605. https://doi.org/10.1126/science.1261605

- Edgcomb, V. (2016). Marine protist associations and environmental impacts across trophic levels in the twilight zone and below. *Current Opinion in Microbiology*, *31*, 169–175. https://doi.org/10.1016/j.mib.2016.04.001

- Elbrächter, M., Schnepf, E., & Balzer, I. (1996). Hemistasia phaeocysticola (Scherffel) comb. nov., Redescription of a Free-living, Marine, Phagotrophic Kinetoplastid Flagellate. *Archiv Für Protistenkunde*, *147*(2), 125–136. https://doi.org/10.1016/s0003-9365(96)80028-5

- Flegontova, O., Flegontov, P., Malviya, S., Audic, S., Wincker, P., de Vargas, C., Bowler, C., Lukeš, J., & Horák, A. (2016). Extreme Diversity of Diplonemid Eukaryotes in the Ocean. *Current Biology*, *26*(22), 3060–3065. https://doi.org/10.1016/j.cub.2016.09.031

- Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., Lieber, M., MacManes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C. N., . . . Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols*, *8*(8), 1494–1512. https://doi.org/10.1038/nprot.2013.084

- Hansen, B., Bjornsen, P. K., & Hansen, P. J. (1994). The size ratio between planktonic predators and their prey. *Limnology and Oceanography*, *39*(2), 395–403. https://doi.org/10.4319/lo.1994.39.2.0395

- Harker, M., Hirschberg, J., & Oren, A. (1998). Paracoccus marcusii sp. nov., an orange Gram-negative coccus. *International Journal of Systematic Bacteriology*, *48*(2), 543–548. https://doi.org/10.1099/00207713-48-2-543

- Hunter, S., Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Das, U., Daugherty, L., Duquenne, L., Finn, R. D., Gough, J., Haft, D., Hulo, N., Kahn, D., Kelly, E., Laugraud, A., Letunic, I., Lonsdale, D., . . . Yeats, C. (2009). InterPro: the integrative protein signature database. *Nucleic Acids Research*, *37*(Database), D211–D215. https://doi.org/10.1093/nar/gkn785

- Khatri, P., Draghici, S., Ostermeier, G., & Krawetz, S. A. (2002). Profiling Gene Expression Using Onto-Express. *Genomics*, *79*(2), 266–270. https://doi.org/10.1006/geno.2002.6698

- Lalli, C., & Parsons, T. R. (1997). *Biological Oceanography: An Introduction* (2nd ed.). Butterworth-Heinemann.

- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

- Logares, R., Audic, S., Santini, S., Pernice, M. C., de Vargas, C., & Massana, R. (2012). Diversity patterns and activity of uncultured marine heterotrophic flagellates unveiled with pyrosequencing. *The ISME Journal*, *6*(10), 1823–1833. https://doi.org/10.1038/ismej.2012.36

- Lukeš, J., Flegontova, O., & Horák, A. (2015). Diplonemids. *Current Biology*, *25*(16), R702–R704. https://doi.org/10.1016/j.cub.2015.04.052

- Lukeš, J., Wheeler, R., Jirsová, D., David, V., & Archibald, J. M. (2018a). Massive mitochondrial DNA content in diplonemid and kinetoplastid protists. *IUBMB Life*, *70*(12), 1267–1274. https://doi.org/10.1002/iub.1894

- Massana, R. (2011b). Eukaryotic Picoplankton in Surface Oceans. *Annual Review of Microbiology*, *65*(1), 91–110. https://doi.org/10.1146/annurev-micro-090110-102903

- Massana, R., Labarre, A., López-Escardó, D., Obiol, A., Bucchini, F., Hackl, T., Fischer, M. G., Vandepoele, K., Tikhonenkov, D. V., Husnik, F., & Keeling, P. J. (2020). Gene expression during bacterivorous growth of a widespread marine heterotrophic flagellate. *The ISME Journal*, *15*(1), 154–167. https://doi.org/10.1038/s41396-020-00770-4

- McCarthy, D. J., Chen, Y., & Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research*, *40*(10), 4288–4297. https://doi.org/10.1093/nar/gks042

- Mukherjee, I., Hodoki, Y., Okazaki, Y., Fujinaga, S., Ohbayashi, K., & Nakano, S. I. (2019). Widespread Dominance of Kinetoplastids and Unexpected Presence of Diplonemids in Deep Freshwater Lakes. *Frontiers in Microbiology*, *10*, 1–13. https://doi.org/10.3389/fmicb.2019.02375

- OmicsBox (2020). OmicsBox – Bioinformatics Solutions. BioBam Bioinformatics. Available online at: https://www.biobam.com/omicsbox/

- Pachiadaki, M. G., Taylor, C., Oikonomou, A., Yakimov, M. M., Stoeck, T., & Edgcomb, V. (2016). In situ grazing experiments apply new technology to gain insights into deep-sea microbial food webs. *Deep Sea Research Part II: Topical Studies in Oceanography*, *129*, 223–231. https://doi.org/10.1016/j.dsr2.2014.10.019

- Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, *14*(4), 417–419. https://doi.org/10.1038/nmeth.4197

- Pernice, M. C., Forn, I., Gomes, A., Lara, E., Alonso-Sáez, L., Arrieta, J. M., del Carmen Garcia, F., Hernando-Morales, V., MacKenzie, R., Mestre, M., Sintes, E., Teira, E., Valencia, J., Varela, M. M., Vaqué, D., Duarte, C. M., Gasol, J. M., & Massana, R. (2014). Global abundance of planktonic heterotrophic protists in the deep ocean. *The ISME Journal*, *9*(3), 782–792. https://doi.org/10.1038/ismej.2014.168

- Powell, S., Szklarczyk, D., Trachana, K., Roth, A., Kuhn, M., Muller, J., Arnold, R., Rattei, T., Letunic, I., Doerks, T., Jensen, L. J., von Mering, C., & Bork, P. (2011). eggNOG v3.0: orthologous groups covering 1133 organisms at 41 different taxonomic ranges. *Nucleic Acids Research*, *40*(D1), D284–D289. https://doi.org/10.1093/nar/gkr1060

- PrabhuDas, M. R., Baldwin, C. L., Bollyky, P. L., Bowdish, D. M. E., Drickamer, K., Febbraio, M., Herz, J., Kobzik, L., Krieger, M., Loike, J., McVicker, B., Means, T. K., Moestrup, S. K., Post, S. R., Sawamura, T., Silverstein, S., Speth, R. C., Telfer, J. C., Thiele, G. M., . . . El Khoury, J. (2017). A Consensus Definitive Classification of Scavenger Receptors and Their Roles in Health and Disease. *The Journal of Immunology*, *198*(10), 3775–3789. https://doi.org/10.4049/jimmunol.1700373

- Prokopchuk, G., Tashyreva, D., Yabuki, A., Horák, A., Masařová, P., & Lukeš, J. (2019). Morphological, Ultrastructural, Motility and Evolutionary Characterization of Two New Hemistasiidae Species. *Protist*, *170*(3), 259–282. https://doi.org/10.1016/j.protis.2019.04.001

- Rocke, E., Pachiadaki, M. G., Cobban, A., Kujawinski, E. B., & Edgcomb, V. P. (2015). Protist Community Grazing on Prokaryotic Prey in Deep Ocean Water Masses. *PLOS ONE*, *10*(4), e0124505. https://doi.org/10.1371/journal.pone.0124505

- Schmitt, R. W. (1995b). The ocean component of the global water cycle. *Reviews of Geophysics*, *33*(S2), 1395–1409. https://doi.org/10.1029/95rg00184

- Schnepf, E. (1994). Light and Electron Microscopical Observations in Rhynchopus coscinodiscivorus spec. nov., a Colorless, Phagotrophic Euglenozoon with Concealed Flagella). *Archiv Für Protistenkunde*, *144*(1), 63–74. https://doi.org/10.1016/s0003-9365(11)80225-3

- Sheldon, R. W., Sutcliffe Jr., W. H., & Paranjape, M. A. (1977). Structure of Pelagic Food Chain and Relationship Between Plankton and Fish Production. *Journal of the*

Fisheries Research Board of Canada, *34*(12), 2344–2353. https://doi.org/10.1139/f77-314

- Sherr, E. B., & Sherr, B. F. (1994). Bacterivory and herbivory: Key roles of phagotrophic protists in pelagic food webs. *Microbial Ecology*, *28*(2), 223–235. https://doi.org/10.1007/bf00166812

- Skovgaard, A. S. (2014). Dirty Tricks in the Plankton: Diversity and Role of Marine Parasitic Protists. *Acta Protozoologica*, *53*(1), 51. https://doi.org/10.4467/16890027AP.14.006.1443

- Stoecker, D., & Pierson, J. (2019). Predation on protozoa: its importance to zooplankton revisited. *Journal of Plankton Research*, *41*(4), 367–373. https://doi.org/10.1093/plankt/fbz027

- Tashyreva, D., Prokopchuk, G., Yabuki, A., Kaur, B., Faktorová, D., Votýpka, J., Kusaka, C., Fujikura, K., Shiratori, T., Ishida, K. I., Horák, A., & Lukeš, J. (2018a). Phylogeny and Morphology of New Diplonemids from Japan. *Protist*, *169*(2), 158–179. https://doi.org/10.1016/j.protis.2018.02.001

- Valach, M., Moreira, S., Hoffmann, S., Stadler, P. F., & Burger, G. (2017). Keeping it complicated: Mitochondrial genome plasticity across diplonemids. *Scientific Reports*, *7*(1), 1–12. https://doi.org/10.1038/s41598-017-14286-z

- Vesteg, M., Hadariová, L., Horváth, A., Estraño, C. E., Schwartzbach, S. D., & Krajčovič, J. (2019b). Comparative molecular cell biology of phototrophic euglenids and parasitic trypanosomatids sheds light on the ancestor of Euglenozoa. *Biological Reviews*, *94*(5), 1701–1721. https://doi.org/10.1111/brv.12523

- Worden, A. Z., Follows, M. J., Giovannoni, S. J., Wilken, S., Zimmerman, A. E., & Keeling, P. J. (2015). Rethinking the marine carbon cycle: Factoring in the multifarious lifestyles of microbes. *Science*, *347*(6223), 1257594. https://doi.org/10.1126/science.1257594