Linking protistan community shifts along salinity gradients with cellular haloadaptation strategies

---

Vom Fachbereich Biologie

der Technischen Universität Kaiserslautern

zur Verleihung des akademischen Grades

„Doktor der Naturwissenschaften" (Dr. rer. nat.)

genehmigte Dissertation,

vorgelegt von M. Sc. Steffen Kühner

Kaiserslautern, 2019

Datum der wissenschaftlichen Aussprache: 1. Juli 2019

Vorsitzender der Prüfungskommission: Herr Prof. Dr. Stefan Kins

Erste Berichterstatterin: Frau Jun. Prof. Dr. Sabine Filker

Zweiter Berichterstatter: Herr Prof. Dr. Matthias Hahn

D 386

Declaration

I, Steffen Kühner, hereby declare that I have produced this present dissertation independently as my own unaided work. The use of other sources or auxiliary material has been properly and fully acknowledged. Furthermore, I declare that this present dissertation has not been submitted complete or in parts to any other institution or university with the intent to obtain an academic degree. Also, I submitted this present dissertation not previously to another governmental or scientific examination board. Additionally, a published paper based on some results of this dissertation can be found under doi 10.1111/1462-2920.14502.

In particular, protected trade names and trademarks are not marked in any way in this work and therefore it cannot be concluded that these are free names or trademarks.

Steffen Kühner

Kaiserslautern, June the ...12...th 2019 _____

Contents

List of figures

List of tables

List of abbreviations

bp	→	base pair
BSA	→	bovine serum albumin
°C	→	temperature in degree Celsius
Ca^{2+}	→	calcium ion
Cl^-	→	chloride ion
cpm	→	counts-per-million
DEA	→	differential expression analysis
dNTPs	→	deoxynucleotides (namely dATP, dGTP, dCTP, dTTP)
Ect	→	ectoine
eDNA	→	environmental DNA
e. g.	→	Latin 'exempli gratia'; for example
ER	→	endoplasmic reticulum
EtBr	→	ethidium bromide
g	→	G force (gravity)
g	→	gram
GB	→	glycine betaine
Gln	→	amino acid glutamine
HOG pathway	→	hyper-osmolar glycerol pathway
^{1}H-NMR spectroscopy	→	proton nuclear magnetic resonance spectroscopy
H_2O_{dest}	→	distilled water
Hz	→	Hertz
I_2	→	iodine
i. e.	→	Latin 'id est'; that is to say
IPC	→	isoelectric point calculator
K^+	→	potassium ion
KEGG	→	Kyoto Encyclopedia of Genes and Genomes
KI	→	potassium iodine
l	→	litre
M	→	Molar (1 mol/l)

MAC	→	macronucleus
mg/l	→	milligram per litre
MIC	→	microneucleus
min	→	minutes
ml	→	millilitre
mRNA	→	messenger-RNA
Na^+	→	sodium ion
NCBI	→	National Center for Biotechnology Information
ng	→	nanogram
NMDS	→	non-metric multidimensional scaling
OG	→	orthologous group
ORF	→	open reading frame
OTU	→	operational taxonomic unit
PCR	→	polymerase chain reaction
qRT-PCR	→	quantitative real-time PCR
ROS	→	reactive oxygen species (e. g. H_2O_2)
rpm	→	rounds per minute
rRNA	→	ribosomal-RNA
RT	→	room temperature
sec	→	seconds
TCDB	→	transporter classification database
TMM	→	trimmed mean of M values
tRNA	→	transfer-RNA
U	→	units
μl	→	10^{-6} l
μmol	→	10^{-6} mol
UPGMA	→	Unweighted Pair Group Method with Arithmetic mean (UPGMA) is a variant of hierarchical cluster analysis

I.   Introduction

1.   Salinity is a structuring factor for organisms

Species are not equally distributed on our planet (Lomolino et al., 2010; Tittensor et al., 2010; Lomolino et al., 2017; Lentendu et al., 2018) and together with biotic interactions, abiotic factors play a fundamental role in the explanation of occurrence or absence of a species (Dunson and Travis, 1991; Benton, 2009; Wingfield and Boonstra, 2013). In each habitat, different abiotic factors are relevant for species distribution. For example, CO_2 concentration, light intensity, nutrient availability, temperature and water movements are limiting factors for aquatic plants (Bornette and Puijalon, 2011). In contrast, cephalopod eggs (*Sepia officinalis*) are sensitive to CO_2 concentration or pH (Gutowska and Melzner, 2009). Besides predation, O_2 concentration is a structuring factor for benthic invertebrate communities (Kolar and Rahel, 1993). O_2 concentrations as well as temperature are important determinants for fishes (Martínez-Álvarez et al., 2005) and the distribution of vertebrates could be structured by pH, high pressure or intense solar radiation (UV-radiation) (Wingfield and Boonstra, 2013). Lozupone and Knight (2007) concluded for bacteria that salinity is the major environmental determinant outcompeting any other physical or chemical factor. Similar results were reported from Wells (1961) regarding oyster beds. Additionally, investigations of macroinvertebrates along a salinity gradient from Piscart et al. (2005) demonstrated that species composition is drastically affected by rising salinity. Also Forster et al. (2012) reported that salinity performed a high selection pressure on ciliates contributing to their community structure, composition and distribution. Furthermore, salinity is the dominant environmental factor regulating aquatic community structures (Witman and Grange, 1998; Verschuren et al., 2000; Waterkeyn et al., 2008; Filker et al., 2017).

An ecotone is defined as transition zone between two biomes and the most investigated ecotone with a distinctive salinity gradient are estuarine systems. Estuaries are mainly located at river deltas reflecting transition zones from freshwater to marine habitats. They are inhabited by former freshwater as well as marine species and only a small proportion of organisms is endemic to estuaries (Gunter, 1961; Crump et al., 2004; Mead et al., 2011). For example, salinity variations had a great impact on benthic faunal structures and biomass in the south-eastern Kattegat on the Swedish west coast (Rosenberg et al., 1992). And for benthic cyanobacteria and diatoms, Nübel et al. (2000) found a clear shift in the community structures below 110 ‰ and above 140 ‰ salinity

indicating that physiological reasons are not negligible as Sorensen et al. (2004) verified for cyanobacteria, anoxygenic phototrophs, sulfate reducers and methanogens of benthic microbial communities in solar salterns of Eilat in Israel. Salinity plays a dominant role in case of species distribution and community structure not only for benthic but also for planktonic organisms (Wu et al., 2006; Wang et al., 2011; Campbell and Kirchman, 2013; Dupont et al., 2014; Filker et al., 2017).

2.   Salinity-dependent habitat classification

The Earth surface is to 71 % covered with water (Perlman and U.S. Geological Survey). Based on salinity, the water masses can be classified as freshwater (< 0.5 ‰ salinity), brackish (0.5 - < 30 ‰ salinity), marine (30 – 40 ‰ salinity) and hypersaline waters (> 40 ‰ salinity) (Venice system 1958). The largest water bodies are marine oceans, which hold about 96.5 % of all Earth´s water. Freshwater is restricted to rivers, lakes, ice caps, glaciers or in the groundwater representing almost 2.5 % of total Earth`s water (Perlman and U.S. Geological Survey). Brackish waters, which are again sub-divided in 'oligohaline' (0.5 – 5 ‰), 'mesohaline' (5 – 18 ‰) and 'polyhaline' (18 – 30 ‰) (Venice system 1958), are mostly transition zones between freshwater and marine systems occurring mainly in river deltas and estuaries. One of the largest and thus most stable brackish water bodies is the Baltic Sea. Because of the freshwater input from rivers of the surrounding countries in the east and only a small connection to the North Sea through the Kattegat and Skagerrak in the west, the Baltic Sea has a well-defined salinity gradient ranging on the surface from 18 ‰ down to less than 5 ‰ from west to east (Fig. 1).
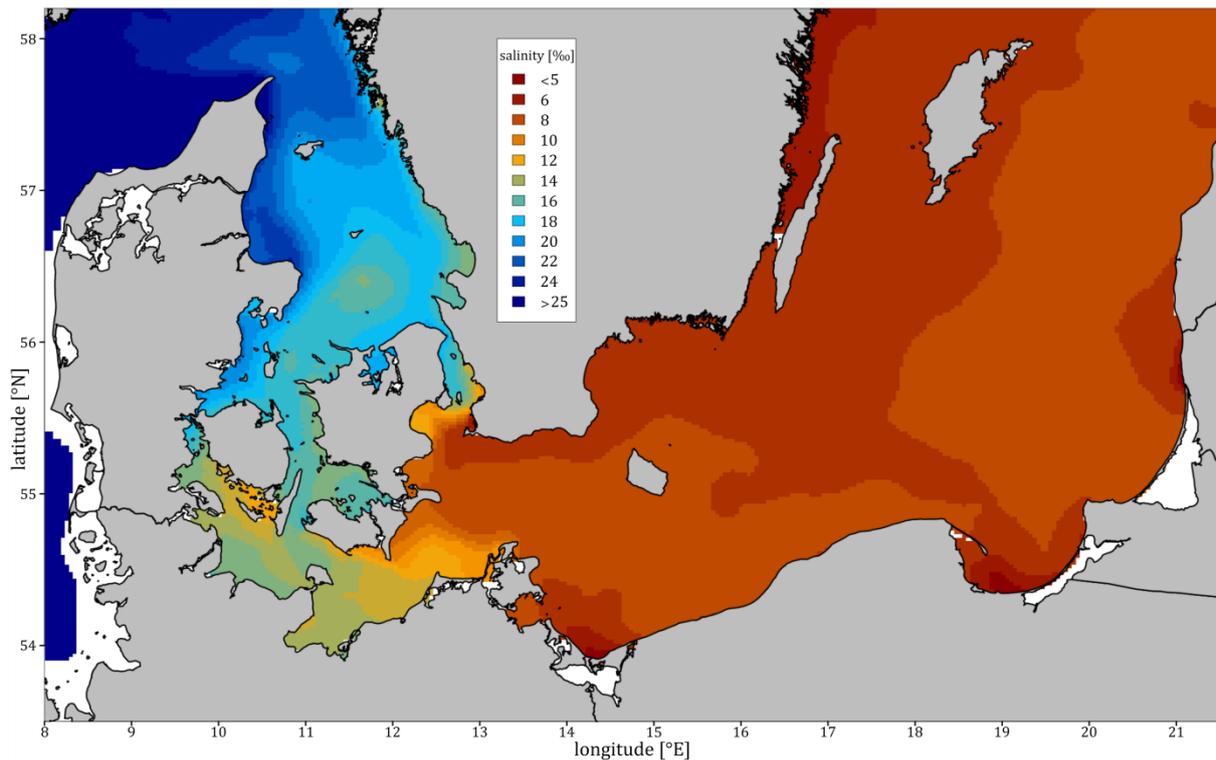
Figure 1: Map highlighting the surface salinity gradient of the Baltic Sea and the neighbouring parts of the North Sea ranging from values greater than 25 ‰ salinity in the west to less than 5 ‰ salinity in the east. The map was kindly provided by Jan Dierking (Geomar Kiel).

Hypersaline habitats are sporadically distributed on each continent around the world (Javor, 1989) and they are either natural or artificial environments. Examples for natural environments are the Great Salt Lake in Utah (USA), which is one of the saltiest water bodies in the world (ca. 290 ‰ salinity; Stephens, 1990) (Fig. 2A), or the Salar de Uyuni in Bolivia, which is the world's largest salt flat (10,000 km^2; Bills et al., 2007) (Fig. 2B). Artificial environments can be found mainly along tropical and subtropical coasts as manmade multi-pond solar systems for commercial salt production like the Ses Salines salterns in Spain (Fig. 2C). Besides their origin, hypersaline environments can also be divided in 'thalassohaline' (Greek: thalassos = Sea) or 'athalassohaline' (Bickel-Sandkötter et al., 1995; Oren, 2006) according to their ion composition in comparison to seawater. 'Athalassohaline' environments are characterized in contrast to seawater by high magnesium and calcium concentrations and low concentrations of sodium, potassium and sulfate (Oren, 2006).

Figure 2: Examples for natural (**A, B**) and manmade (**C**) hypersaline environments. **A**: The Great Salt Lake (Utah, USA); **B**: The Salar de Uyuni (Bolivia); **C**: Overview of the different salt ponds of the Ses Salines salterns (Ibiza, Spain) as an example for a manmade multi-pond solar system to produce salt. The different colours of the certain ponds indicate different ecosystem phenotypes by reflecting dense blooms of halophilic organisms.[1]

Investigating the species and species compositions in these hypersaline environments is of outstanding importance for science. Hypersaline environments are one of the least understood ecosystems on Earth resulting in a frequent discovery of novel halophilic organisms (Filker et al., 2015). Natural changes like global warming and anthropogenic exploitation of water resources (e. g. Aral Sea; Micklin, 2007), which is based on overfertilization due to increasing population, will boost the development and

---

[1] **A**: https://c1.staticflickr.com/5/4400/37106367906_c348872d9d_b.jpg; **B**: https://cdn.moja-travel.net/image/resize/1200x-/e601232c-7a41-4334-8417-3ea34c2fdcaf-188b00005-salar-de-uyuni-kompakt-map.jpg; **C**: http://www.traveltoibiza.es/wp-content/uploads/2015/02/Salinas-4-FU-VM-608x367.jpg

formation of hypersaline environments in the future. Salinisation and an increased desertification are serious threads for human survival because nearly 2.4 billion people or 40 % of the human population (United Nations, 2017) live in coastal and 2.1 billion people in semi-arid or arid areas (United Nations) with limited freshwater resources. Coastal desalinisation plants for freshwater generation and solar salterns for salt production are technical opposed inventions to satisfy human needs for water and salt. But despite the human ingenuity, mankind is reliant on functioning ecosystem services including for instance the supply of food or fresh air. Therefore, a full understanding of ecosystems is essential to be able to make reasonable predictions about the impact of alternating environmental conditions caused by i. e. climate change. To understand the ecosystems of hypersaline environments, a fundamental knowledge about species, their distribution and metabolic adaptation strategies is essential (Chapin et al., 1997).

3.   Biodiversity of protists in salinity-dependent environments

Protists strongly influence salinity-dependent ecosystems because protists are essential components of the microbial loop and hold key positions in pelagic food webs (Barber, 2007). Although protists are the most abundant and diverse eukaryotes on our planet occurring nearly in each ecosystem (Patterson, 1999; Baldauf, 2008), most described species and large protistan diversity studies originated almost exclusively from marine environments and hypersaline habitats were mostly irgnored in this respect (Elloumi et al., 2009; Heidelberg et al., 2013; Filker et al., 2015; Filker et al., 2017).

For the Baltic Sea as the largest semi-enclosed brackish water body, Telesh and colluegues (2011; 2013) conducted a meta-analysis of mainly microscopic phyto- and zooplankton data discovering protistan species maximum and high plankton diversity in the horohalinicum (5 – 8 ‰ salinity) of the Baltic Sea. But this inverse trend comparing to macroorganisms is still debatable because of the absence of a large scale diversity study based on high-throughput sequencing data. Technical biases and the insufficiency of microscopy compared to molecular diversity are well-known in the scientific community and were discussed by several authors (Bachy et al., 2013; Stoeck et al., 2014; Visco et al., 2015; Rivera et al., 2018). Additionally, Logares et al. (2009) reported that marine–freshwater transitions had been infrequent events during the

diversification of microbes and Oren (2001) justified the limit of environmental transitions with the evolution of adaptation mechanisms and the different energetic costs for osmoregulation to live in such environments, like brackish waters as transition zone for marine and freshwater species. But the microbial advantages like large population size (Giovannoni and Stingl, 2005), ability for long-distance passive dispersal (Kellogg and Griffin, 2006) and high reproduction rate (Weisse, 2008; Logares et al., 2009) may allow protists to cross this marine-freshwater salinity barrier. Therefore, only high-throughput sequencing studies could highlight the real protistan diversity of brackish environments by catching and identifying all occurring freshwater, brackish as well as marine taxa. Concluding from the absence of experimental as well as uniform high-throughput sequencing datasets, currently relying only on taxonomic surveys (Logares et al., 2009), the question about the existence of a protistan marine–freshwater transition boundary is still unsolved and it is unclear whether salinity could be a structuring factor for protistan communities in low-salt environments, such as the Baltic Sea.

When it comes to hypersaline environments, the general view was that those environments are habitats with a reduced microbial species diversity (Reed et al., 1984). This is mainly due to the high osmotic pressure caused by high salt concentrations. Congruently, Heidelberg et al. (2013) reported that hypersaline systems are dominated by prokaryotes and the communities are controlled by haloarchaea with a low eukaryotic diversity (Oren, 2002a). Cantrell et al. (2006) and Oren (1999a) argued eukaryotic 'extremophiles' can only live under high salt concentrations as long as enough energy resources are available to maintain their metabolic cell functionality. In contrast, diverse unicellular eukaryotes were reported from hypersaline waters a long time ago (Butschinsky, 1897; Kirby, 1932; Ruinen, 1938; Harding and Simpson, 2018). Until now, quantitative data on eukaryotic community assemblages is still limited (Heidelberg et al., 2013) because most research in hypersaline environments focused for decades on archaea and bacteria with neglecting microbial eukaryotes (Oren, 2002a). For example, Elloumi et al. (2009) found indeed marked changes in community compositions and biomass along a salinity gradient, especially at 150 ‰ and 350 ‰ salinity, but only a few species are detected as typical halophiles, namely *Dunaliella salina*, *Fabrea salina* and *Artemia salina*. A few decades ago, the following view was widely accepted for hypersaline ecosystems: Chlorophyte algae of the genus *Dunaliella*

were the main primary producer (Oren, 2005), which were consumed by the brine shrimp *Artemia salina* as the main predatory organism (Pedrós-Alió et al., 2000). But recently, combining studies using growth experiments of salt-tolerant organisms, microscopy, environmental sequencing, molecular phylogenetics and transcriptomics greatly improved the scientific knowledge of microbial eukaryotes sampled in hypersaline environments (Triadó-Margarit and Casamayor, 2013; Harding and Simpson, 2018). As a result, at least one representative in almost each eukaryotic 'supergroup' has been reported for hypersaline environments (Hauer and Rogerson, 2005; Harding and Simpson, 2018). Furthermore, the list of protists with hypersaline origin is still expanding since high-throughput sequencing and other molecular techniques provide a continuous flow of unknown sequences as potential source for discovering novel species (Epstein and López-García, 2008; Filker et al., 2015). To the present state of knowledge, heterotrophic ciliates, amoebae and flagellates are the dominant protistan groups in hypersaline environments (Hauer and Rogerson, 2005) due to their dominant role in food web structures and biogeochemical cycles (Oren, 2002a; Pedrós-Alió, 2005; Elloumi et al., 2009).

Interestingly, several authors investigated species richness in multi-pond solar salterns along a wide range of different salinities and reported a notable decrease of protistan species richness with increasing salinity (Hauer and Rogerson, 2005; Filker et al., 2017). Nonetheless, Filker et al. (2015) used a deep sequencing approach to detect novel protistan plankton diversity in contrast to Elloumi et al. (2009) using optical techniques. Consequently, Filker et al. (2015) found a higher diversity in solar saltern ponds than Elloumi et al. (2009) and other microscopy based studies. To eliminate potential biases due to geographical differences, seasonal effects or the occurrence of cosmopolitans, like the green algae *Dunaliella salina* or the ciliate *Fabrea salina*, the large-scale sequencing study of Filker et al. (2017) investigated protistan community structures in different solar saltern systems from Europe (France, Spain, Portugal) and South America (Argentina, Chile) with salinities ranging from 39 ‰ to 440 ‰. The most novel genetic diversity was detected at 270 ‰ salinity. The highest degree of novelty was found within the ciliophora at intermediate salinities (140 – 240 ‰) even though they are considered as the best known taxon group in the kingdom Protista (Filker et al., 2015). Interestingly, this unprecedented view into protistan communities from hypersaline environments showed that salinity appeared as stronger selection factor

than geography. Furthermore, four salinity classes were defined with boundaries for protistan species turnover in between, namely 40 – 90 ‰, 140 – 240 ‰, 270 – 360 ‰ and 380 – 440 ‰ (Fig. 3). Additionally, the transition boundary with the strongest selection was established between 240 ‰ and 270 ‰ salinity. At this point, the separation of halotolerant from halophile protists was assumed because halophile organisms are only able to grow with a certain minimum of salt (DasSarma and Arora, 2002). Haloterant species, instead, are able to grow under low-salt conditions and can tolerate higher salt concentrations than they need for growth (Kushner, 1978). However, the metabolic properties remain unknown, which may be responsible for protistan community shifts and their observed difference in taxonomic membership and degree of complexity. Filker et al. (2017) explained these findings with the assumption that different evolutionary lineages of protists had evolved distinct haloadaptation strategies requiring different cellular responses to master their life with high salt concentrations.
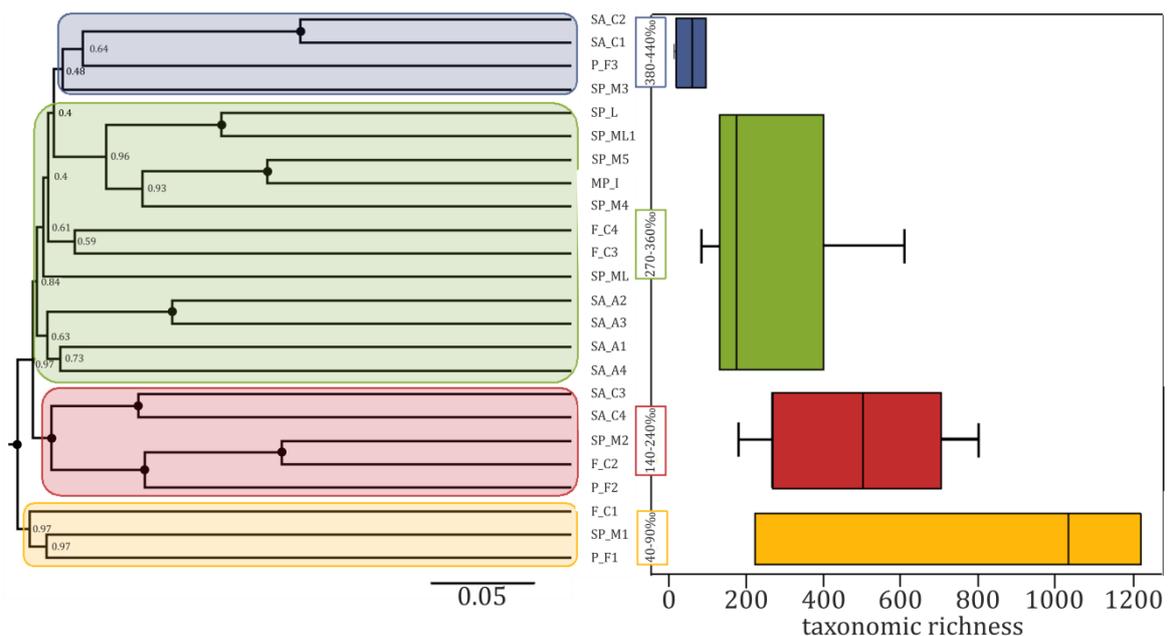
Figure 3: Salinity is one of the major abiotic determinants for shifting protistan plankton communities. The diversity analysis as UPGMA on the left hand side is based on binary Sørensen-Dice distances and demonstrates a clear structuring pattern according to salinity. Bootstrap values are shown as indicator for strength of relationships in the dendrogram (• = 1). Four salinity classes are implied, namely 40 – 90 ‰, 140 – 240 ‰, 270 – 360 ‰ and 380 – 440 ‰, and their taxonomic richness declines with increasing salinity (right hand side; results are based on one-way ANOVA with alpha = 0.05, p<0.01). The data was generated by targeting the V4 region of the SSU rRNA. Figure was slightly modified from Filker et al. (2017) by courtesy of Sabine Filker (University of Kaiserslautern).

4.   Known haloadaptation strategies in protists

Until recently, almost nothing was known about potential haloadaptation strategies in protistan plankton as only a handful of studies focused on this topic (Ben-Amotz and Avron, 1980; Lages et al., 1999; Petrovicˇ et al., 2002; Kogej et al., 2005; Gunde-Cimerman and Plemenitaš, 2006). Knowledge about haloadaptation strategies was gained for decades almost exclusively from bacteria, archaea, and a few fungi. Two main fundamentally different cellular strategies have been identified for osmotic adjustment (Santos and da Costa, 2002): the 'high-salt-in' strategy (Joghee and Jayaraman, 2014) and the 'low-salt – organic-solutes-in' strategy (Brown, 1976, 1990; Welsh, 2000; Santos and da Costa, 2002).

The 'high-salt-in' strategy is based on the selective influx and accumulation of inorganic ions, mostly chloride (Cl^-), potassium (K^+) or sodium (Na^+) ions, to reach osmotic equilibrium (Galinski and Trüper, 1994; Santos and da Costa, 2002; Joghee and Jayaraman, 2014). Until now, only extreme halophile archaea, e. g. *Halobactrium salinarum* and *Haloferax volcanii* (Soppa, 2006) and some halotolerant bacteria were described using this 'high-salt-in' strategy (Blumwald et al., 1983; Oren, 2002b, 2008). Protistan representatives for that strategy are still missing. Depending on the extracellular salinity, intracellular ion concentrations can achieve more than 5 M (Christian and Waltho, 1962; Ginzburg et al., 1970) and normally, such high intracellular ion concentrations are very harmful to macromolecules and the enzyme machinery of the cell (Lanyi, 1974). Thus, the complete enzyme machinery of 'high-salt-in' strategists has to be adapted to function under such conditions. Therefore, typically proteomes of 'high-salt-in' strategists are highly acidic by presenting more negatively than positively charged protein residues on the protein surface (Paul et al., 2008; Harding et al., 2016). For example in the extreme halophilic bacterium *Salinibacter ruber*, this excess of negative charge is balanced by water molecules attached to the protein surface increasing protein solubility as well as preventing their proteins to aggregate at high salinities (Madern et al., 2000; Oren et al., 2002; Oren, 2013). As a consequence, without an adequate salinity in their surrounding cytoplasm halophilic proteins will denature and lose their function (Lanyi, 1974; Danson and Hough, 1997). Due to this highly specialized protein and enzyme machinery, 'high-salt-in' organisms are mostly restricted to a relatively limited range of high-salt concentrations and lack the flexibility to adapt to great changes in external salinity.

A completely different haloadaptation strategy is used if an organism is unable to tolerate high intracellular amounts of inorganic ions by a well-adapted proteome, e. g. because of high energy costs for genetic or post-transcriptional or enzyme modifications (Oren, 1999a). This opposite strategy is called 'low-salt – organic-solutes-in' strategy and is based on the accumulation of so called 'compatible solutes', which are highly water-soluble, low molecular weight, organic compounds (Joghee and Jayaraman, 2014). They are osmotically active and normally do not interfere with cellular functions (Brown, 1976; Zhao, 2005). Compatible solutes can be accumulated through *de novo*-synthesis (Galinski and Trüper, 1982) or via uptake by transporter from the environment if they are present (Santos and da Costa, 2002). They counteract the water efflux under hypertonic growth conditions (Kempf and Bremer, 1998). Besides their predominant osmotic function, compatible solutes stabilise biomolecules (Santos and da Costa, 2002; Joghee and Jayaraman, 2014), and therefore they could also play a protective function against high temperature, freezing, desiccation and other denaturing conditions (Welsh, 2000). The utilisation and accumulation of compatible solutes were investigated firstly in prokaryotes (Empadinhas and da Costa, 2008), especially bacteria (Imhoff, 1986), but meanwhile several eukaryotes were described using the 'low-salt – organic-solutes-in' strategy. The most studied species in this respect is the algae *Dunaliella salina* and its relatives (Oren, 2005). *Dunaliella salina* accumulates glycerol as compatible solute and therefore, its membrane has to have lower glycerol permeability (Oren, 2005). Nowadays, more than 50 different compatible solutes are known to science, chemically divided into amino acids and their derivatives, sugars and their derivatives or phosphate-containing solutes. Their chemical structure could be very simple (e. g. proline) or complex (e. g. ectoine, di-*myo*-inositol-1,1´-phosphate) with diverse chemical characteristics being zwitterionic, uncharged or anionic (Roberts, 2005; Empadinhas and da Costa, 2008).

In experimental studies based on ^{1}H-NMR spectroscopy, Weinisch et al. (2018b; 2018a) detected glycine betaine and ectoine as compatible solutes in the ciliates *Cyclidium glaucoma*, *Fabrea salina*, *Pseudocohnilembus persalinus*, *Schmidingerothrix salinarum* and a species of the genus *Euplotes*. Together with their results via ion imaging, they concluded *S. salinarum* is a 'low-salt – organic-solutes-in' strategist accumulating glycine betaine and ectoine as compatible solutes (Weinisch et al., 2018b). Such equivalent proof for the usage of hydroxyectoine and *myo*-inositol as compatible

solutes in the two nanoflagellates *Halocafeteria seosinensis* and *Pharyngomonas kirbyi* is still missing in the study of Harding et al. (2016). But nevertheless, transcriptome analyses of these two nanoflagellates, especially *H. seosinensis*, revealed additional cellular reactions for living in high-salt environments, besides the synthesis of compatible solutes. After salt-up shock, *H. seosinensis* showed an increased amount of stress response related transcription factors, cyclic nucleotide-dependent signalling molecules for signal transduction, an overexpression of a Na^+/H^+ antiporter for sodium expulsion, a reduced oxygen respiration, sterol and phosphatidylethanolamine biosynthesis and increased fermentation processes instead (Harding et al., 2017). Additionally, protein damage control and glutathione-dependent reactive oxygen species (ROS) detoxification were important processes indicating stress management as an important aspect of haloadaptation (Harding and Simpson, 2018). Based on their transcriptome, *H. seosinensis* and *P. Kirbyi* were classified as salt-excluders using the 'low-salt – organic-solutes-in' strategy (Harding et al., 2016; Harding et al., 2017). Despite the fact, that many other protists can grow in culture, e. g. *Trimyema koreanum*, *Platynematum salinarum*, *Euplotes qatarensis,* their salt-stress adaptation strategy has never been studied in detail (Harding and Simpson, 2018).

However, apart from protists, some micro-eukaryotic fungi were intensively investigated regarding their haloadaptation mechanisms. The black yeast *Hortaea werneckii*, which uses glycerol as compatible solute (Petrovicˇ et al., 2002), excludes sodium ions via Na^+/H^+ antiporter or ATP-dependent Na^+ pump (Kogej et al., 2005). Additionally, *H. werneckii* synthesizes dihydroxynaphthalene (DHN) melanin (Gunde-Cimerman and Plemenitaš, 2006) and has a lower sterol-to-phospholipid ratio to make their membranes more fluid together with three fatty acid-modifying enzymes (Gunde-Cimerman and Plemenitaš, 2006; Gostincar et al., 2009). Another special behavior was observed for the halophile yeast *Debaryomyces hansenii*. This yeast utilizes apparently parts of both main strategies: Besides the accumulation of glycerol as a compatible solute, additional Na^+ are possibly incorporated at lower salinities. With increasing salinity at the cells outside, those Na^+ are substituted by K^+ (Prista et al., 1997; Prista et al., 2005; Ramos, 2005). Therefore, it can be assumed for micro-eukaryotes that lot of different specifications of the two main strategies ('high-salt-in' or 'low-salt – organic-solutes-in') or even a mixture ('hybrid strategy') may exist and scientists are far away to have a full and complete understanding about haloadaptation strategies, which enable

protists to live in hypersaline environments. Consequently, it can only be hypothesized that community shifts (Fig. 3), which were described by Filker et al. (2017), are based on a changing metabolome of the community resulting from a shift regarding the used haloadaptation strategy.

5. Goals of this study

My PhD project aimed to explain protistan community shifts along salinity gradients through the use of different haloadaptation strategies by protists. To achieve this goal, the following three main subjects were addressed:

Subject I:

eDNA metabarcoding of protistan plankton communities in the Baltic Sea

Protistan diversity studies in hypersaline environments (salinity > 40 ‰) revealed specific salinity boundaries, which seem to select for and prevent the dispersal of some species. In contrast, for low-salt habitats (< 40 ‰), corresponding investigations were either missing or inconclusive due to the data basis. In the first part of my thesis, I thus aimed to *provide reliable data for protistan plankton thriving in low-salt habitats and subsequently to reveal their diversity, community composition and distribution patterns.* I sampled environmental protistan DNA along a 600 km transect in the Baltic Sea, covering salinities between 7.87 – 22.05 ‰ and applied high-throughput next-generation sequencing to analyse obtained protistan eDNA metabarcodes. *By integrating these data with datasets obtained by Filker et al. (2017), I extended the thus far investigated salt gradient in order to infer whether observed salinity boundaries stay valid and whether specific salt transition boundaries for species turnover also occur in low-salt habitats.*

Subject II:

Analyses of gene expression patterns of microeukaryotes along a salinity gradient

The results from Filker et al. (2017) and subject 'I' revealed significant shifts in protistan community composition along low-brackish to extreme-hypersaline conditions occurring at specific salinities. The most likely explanation for the observed

protistan distribution patterns is the requirement of different adaptive mechanisms to cope with high salt concentrations. In the second part, I thus investigated whether this hypothesis holds true and *whether the observed shifts in protistan community composition can be linked to different haloadaptation strategies such as the 'high-salt-in' or the 'low-salt – organic-solutes-in' strategy or any combination of both ('hybrid' strategy).* For this, I conducted environmental, comparative meta-transcriptome analyses of natural protistan communities thriving between 40 and 380 ‰ salinity. This approach enabled for the investigation of potential physiological functions of whole protistan communities in response to salinity without cultivation-biases.

Subject III:

Time-resolved transcriptome analysis of *Schmidingerothrix salinarum*

The findings of subject 'II' pointed to a preferred usage of the 'low-salt – organic-solutes-in' strategy by protists tolerating salinities up to 150 ‰. This strategy, using an experimental approach, was recently also identified in the model ciliate *Schmidingerothrix salinarum*, together with hints that the ciliate is able to synthesise and accumulate the compatible solutes glycine betaine and ectoine (Weinisch et al., 2018b). However, detailed and comprehensive information about involved enzymes, specific pathways and cellular processes associated with the 'low-salt – organic-solutes-in' strategy applied by ciliates (as representatives of protists) could not be obtained with this approach based on ^{1}H-NMR spectroscopy. *I, therefore, conducted a time-resolved transcriptome analysis of this model ciliate to investigate its cellular expression dynamics after osmotic up-shock and during salt acclimation.*

II. Material and Methods

1.   eDNA metabarcoding of protistan plankton communities in the Baltic Sea

In this part of my PhD thesis I obtained protistan eDNA metabarcodes along a 600 km transect in the Baltic Sea and analyses were divided into the chapters '1.1 Sampling', '1.2 DNA extraction, PCR amplification and sequencing', '1.3 Sequence data processing and taxonomic assignment' as well '1.4 Statistical analyses'. The methodology was already published and largly copied from Filker et al. (2019). Equally, with exception of sub-chapter '1.4 Salinity-dependent transition boundaries for protistan species turnover', the associated part of the results ('III. Results: 1. eDNA metabarcoding of protistan plankton communities in the Baltic Sea') was mostly transferred from Filker et al. (2019) and also to some extent the first part of the discussion ('IV. Discussion: 1. eDNA metabarcoding of protistan plankton communities in the Baltic Sea'). However, I integrated my data with datasets obtained by Filker et al. (2017) to extend the thus far investigated salt gradient with low-salt habitats ranging now from 8 to 440 ‰ salinity.

1.1   Sampling

Water samples were taken during research cruise *AL491* with the *R/V* Alkor in the Baltic Sea in April 2017. A rosette of 24 10 l-Niskin-bottles, equipped with a conductivity, temperature, pressure sensor (CTD, ADM Elektronik, Krems II, Germany) and an oxygen probe (AMF, Rostock, Germany) was used to measure the respective environmental factors (Suppl. Tab. 1). Salinities were derived from conductivity values. Seventeen water samples were collected at 6 m depth (surface samples) along a horizontal salinity gradient covering a 600 km transect with salinities ranging between 7.87 ‰ in the east and 22.05 ‰ in the west (Fig. 4, Suppl. Tab. 1).
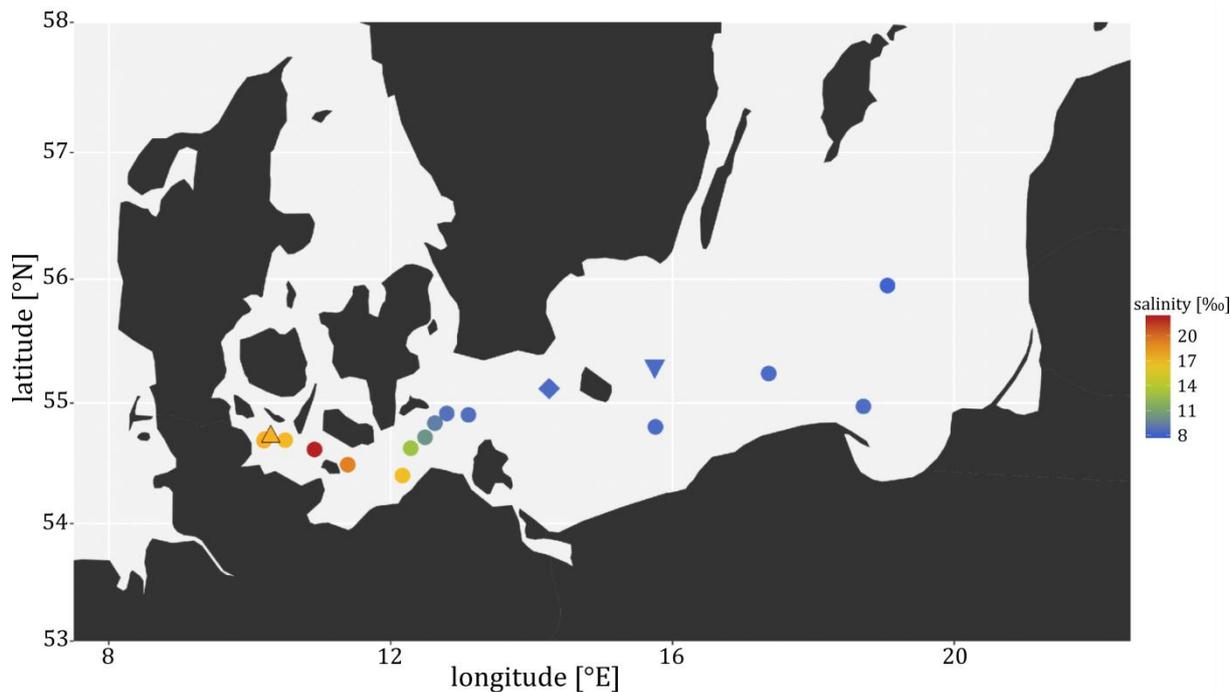
Figure 4: Map of the sampling stations in the Baltic Sea of cruise *AL491* with *RV* Alkor colour-coded according to the surface water salinities (triangle: Kiel Bight; diamond: Arkona Basin; inverted triangle: Bornholm Basin). The map was slightly modified from Filker et al. (2019).

Using a three-head peristaltic pump (Rotarus, Hirschmann, Germany), five litres of water from three different Niskin-bottles per sampling point were filtered through polycarbonate membrane filters with a pore size of 0.65 μm (Durapore, Millipore, Germany) to capture microeukaryotic cells. Triplicate filters were prepared in order to account for a potential sampling bias. Thus, 51 filters (=17 sampling points * 3 replicates) were prepared in total. Each filter was preserved in 3 ml RNA*later* (Qiagen, Hilden, Germany), incubated at 4 °C for 24 h and stored at -20 °C until further processing in the lab.

1.2 DNA extraction, PCR amplification and sequencing

Extraction of total environmental DNA (eDNA) from all triplicate filters per sampling point followed the protocol of Qiagen´s All Prep DNA/RNA Mini Kit (Qiagen, Hilden, Germany) (Suppl. Tab. 2). eDNA from the same sampling points were then pooled and subjected to a spectro-photometrical determination of concentration using a NanoDrop2000 (Thermo Scientific, Wilmington, DE, USA).

Three eDNA replicates (circa 50 ng each) of each sampling point were used as template to amplify the hypervariable V4-region of the SSU rRNA gene in a polymerase chain reaction (PCR). This circa 450 base pair (bp) fragment serves as metabarcode to distinguish eukaryotic species in metacommunity surveys. The barcode was targeted with the primer pair TAReuk454FWD1 (5'- CCAGCA(G/C)C(C/T)GCGGTAATTCC -3', *S. cerevisiae* position 565-584) and TAReukREV3 (5'- ACTTTCGTTCTTGAT(C/T)(A/G)A -3', *S. cerevisiae* position 964-981) as proposed by Stoeck et al. (2010). The PCR master mix consisted of 5x GC buffer (New England Biolabs, Frankfurt am Main, Germany), 20 µg/µl bovine serum albumin (BSA; New England Biolabs, Frankfurt am Main, Germany), 10mM of each deoxynucleotides (dNTPs, namely dATP, dGTP, dCTP, dTTP; New England Biolabs, Frankfurt am Main, Germany), 10 µM forward primer TAReuk454FWD1 (biomers.net GmbH, Ulm, Germany), 10 µM reverse primer TAReukREV3 (biomers.net GmbH, Ulm, Germany), nuclease-free water (Carl Roth GmbH & Co. KG, Karlsruhe, Germany) and 1 U of a Phusion High Fidelity DNA polymerase (New England Biolabs, Frankfurt am Main, Germany), which was used to minimize errors during the elongation. The PCR protocol for amplifying the V4 region covered an initial denaturation at 98 °C for 30 sec, 26 cycles of denaturation (98 °C, 10 sec), annealing (63 °C, 30 sec) and elongation (72 °C, 30 sec) and a final elongation at 72 °C for 10 min (Tab. 1). The protocol was performed on an iCycler thermal cycler (Bio-Rad Laboratories GmbH, Munich, Germany).

Table 1: PCR program for amplifying the hypervariable V4 region of the SSU rRNA

initial denaturation	98 °C	30 sec	
denaturation	98 °C	10 sec	26 x
annealing	63 °C	30 sec	amplification
elongation	72 °C	30 sec	cycles
final elongation	72 °C	10 min	

Afterwards, an agarose gel electrophoresis was done to check the quality of the received PCR products. Therefore, 0.9 g agarose (Carl Roth GmbH, Karlsruhe; Germany) was dissolved in 100 ml TAE buffer (AppliChem GmbH, Darmstadt, Germany) by heating in a microwave for 2 min. After adding of 4.5 µl ethidium bromide (EtBr; Carl Roth GmbH, Karlsruhe; Germany) and hardening of the gel, 10 µl of each PCR product

together with 2 µl loading dye (Qiagen, Hilden, Germany) was loaded into pockets separately next to the 1 kb DNA ladder (New England Biolabs, Frankfurt am Main, Germany). The gel electrophoresis was run for 30 min at 90 V (Suppl. Fig. 1).

Triplicate PCR products per sampling point were pooled prior to sequencing library preparation. Sequencing libraries were constructed using the NEB Next® Ultra™ DNA Library Prep Kit for Illumina (New England Biolabs, Frankfurt am Main, Germany), including a sample specific tagging with 6-bp identifier fragments to allow demultiplexing in downstream bioinformatics data processing. The quality of the 17 sequencing libraries was evaluated with an Agilent Bioanalyzer 2100 system (Agilent Technologies, Waldbronn, Germany) prior to pooling. The final library was sequenced on an Illumina MiSeq platform generating 2 x 250 bp reads by SeqIT GmbH & Co. KG (Kaiserslautern, Germany). The raw sequence data was deposited at the sequence read archive of the National Center for Biotechnology Information (NCBI) under the accession number SRP156873.

1.3 Sequence data processing and taxonomic assignment

Raw paired-end reads were assembled and demultiplexed by SeqIT GmbH & Co. KG using custom scripts. Quality control and filtering of raw reads was done with the 'split_library.py' script as implemented in QIIME v1.8 (Caporaso et al., 2010), keeping only reads with a quality of >Q20, exact primer matches, no ambiguous nucleotides and a minimum length of 300 bp. Since chimera formation frequently occurs during PCRs, processed reads were furthermore controlled with the *de novo* chimera detection option of UCHIME (Edgar et al., 2011) in order to filter out potential chimeric artefacts. High-quality reads were then clustered in operational taxonomic units (OTUs) using SWARM v. 2.0.5 (Mahé et al., 2015) with d = 1 and the fastidious option. To assign taxonomic identities to the OTUs, the most abundant sequence of each OTU was extracted and annotated with BLASTN (Altschul et al., 1990) against NCBI´s non-redundant nucleotide reference database (NCBI-GenBank Flat File Release 220.0). OTUs, that consisted of only one or two reads and occurred only in one sample were regarded as sequencing noise and excluded from downstream analyses, as were all metazoan and embryophyte OTUs. This final OTU table (provided on appended CD-Rom) served as basis for all downstream statistical analyses.

1.4 Statistical analyses

In general unless stated otherwise, all statistical analyses were conducted in R v. 3.3.2 (R Core Team, 2013) using the 'vegan' package (Oksanen et al., 2015) and 'ggplot2' (Wickham, 2016) for graphical representations. All R scripts are provided digital on the appended CD-Rom or upon request.

Sampling curves for species richness (Hill number q=0) with 95 % confidence intervals were calculated with the 'iNext' package (Hsieh et al., 2016) in order to investigate the degree of sample saturation, sample-size based rarefaction and extrapolation. Samples were randomly normalized to the smallest sample size (n = 70,289) to account for uneven sample sizes and enable comparability. The Shannon-based effective number of species (Jost, 2007), the Simpson index and OTU richness served as alpha-diversity parameters for each sample. To assess the influence of environmental factors to the alpha-diversity, the Pearson correlation coefficient was calculated for combinations of the samples alpha-diversity and the environmental factors salinity, temperature and oxygen, respectively. Correlations were considered significant if $p < 0.05$ and $r > \pm 0.75$. Beta diversity, i. e. (dis)similarity between samples, was investigated using the incidence-based Jaccard index. The Jaccard index was chosen to measure species turnover (Tuomisto, 2010) to avoid potential biases based on the different sequence abundances within each sample due to different gene copy numbers in protists (Zhu et al., 2005; Dunthorn et al., 2014). Obtained similarity values were then transformed into a distance matrix serving as basis for an unweighted pair group method with arithmetic mean (UPGMA) and a non-metric multidimensional scaling (NMDS) analysis. The envfit function of the 'vegan' package in R was used to fit environmental vectors to the NMDS ordination. With a Monte Carlo analysis of 1000 permutations, the fit (R^2) of each variable to the ordination was assessed. An analysis of variance (one-way ANOVA) of alpha-diversity values and the environmental parameter between sample groupings was performed using the 'stats' package to investigate significant differences between these groups.

A partial Mantel test was conducted to assess a potential spatial effect on the protistan community composition (i. e. beta diversity). Therefore, geographic distances were determined based on the longitude/latitude data using the 'haversine' formula as implemented in the script provided by Chris Veness at http://www.movable-type.co.uk/scripts/latlong.html (Veness, 2002-2017). The differences between two sample pairs of the three variables, beta diversity, geographic distances and salinity, were correlated with each other and respectively tested.

2. Analyses of microeukaryotic gene expression patterns along a salinity gradient

In this part of my PhD thesis I investigated potential physiological functions of whole protistan communities in response to salinity without cultivation-biases by conducting environmental, comparative meta-transcriptome analyses of natural protistan communities thriving between 40 and 380 ‰ salinity. These meta-transcriptome analyses were divided in field work (see chapter '2.1 Sampling'), lab procedures (see chapters '2.2 RNA extraction and quality control', '2.3 Poly-A enrichment and mRNA sequencing') and sequence data analysis (Fig. 6) containing several bioinformatic analyses and statistics (see chapter '2.4 Bioinformatics and statistics').

2.1 Sampling

Water samples were collected in the solar saltern 'Salinas d`Es Trenc' (Mallorca, Spain) (39°21′04″ N, 3°00′42″ E) (Fig. 5A) in early October 2017 by Dr. Sabine Filker (Dept. of Molecular Ecology, University of Kaiserslautern, Germany). Salinity was measured using a hand-held refractometer (Atago, Tokyo, Japan). In order to cover a pronounced salinity gradient as well as the previously identified salinity classes, samples were taken at the saltern inlet exhibiting salinity of 40 ‰ and from six subsequent ponds with salinities of 110 ‰, 150 ‰; 200 ‰ (Fig. 5B), 240 ‰ (Fig. 5C), 300 ‰ (Fig. 5D) and 380 ‰. For each site, three times three litres of water were filtered through polycarbonate membrane filters with a pore size of 0.65 μm (Durapore, Millipore, Germany) using a peristaltic pump (Rotarus, Hirschmann, Germany) until clogging was observed. Filters were directly preserved in 3 ml RNA*later* (Qiagen, Germany) to stabilize collected DNA and RNA, incubated at 4 °C for 24 h and stored at -80 °C until further processing.
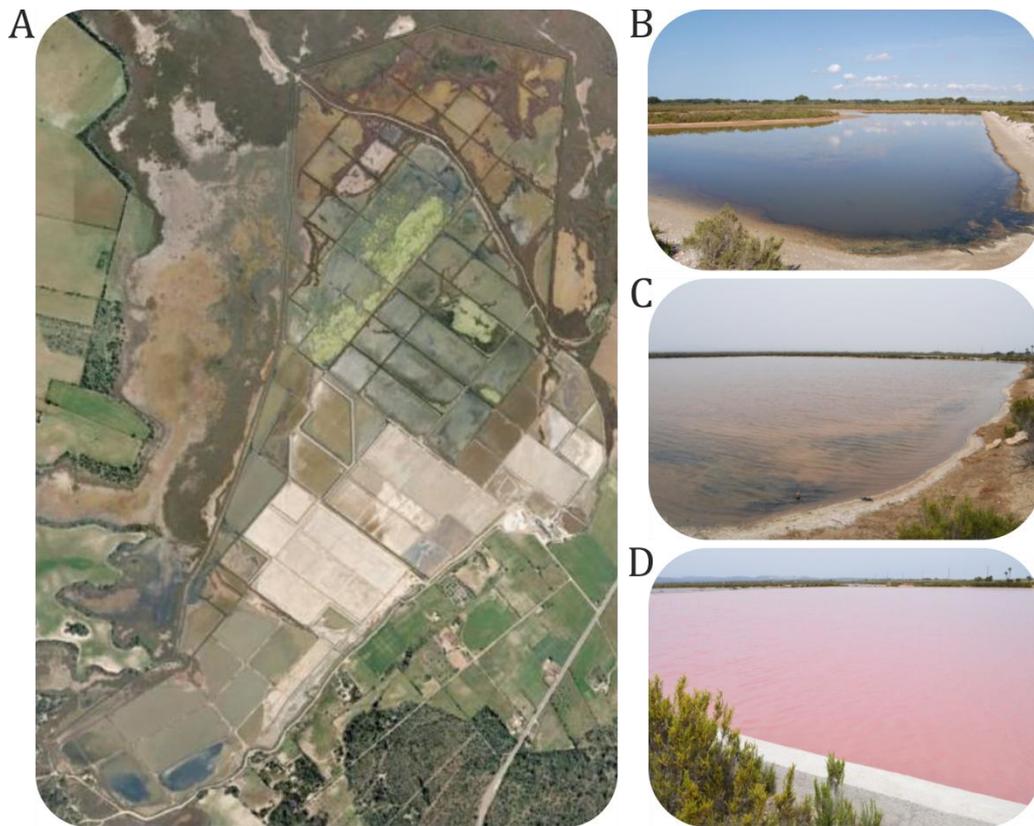
Figure 5: Solar saltern 'Salinas d`Es Trenc' (Mallorca, Spain). **A:** Satellite overview from Google Earth (02/25/2019); **B:** Evaporation pond with 200 ‰ salinity, **C:** Evaporation pond with 240 ‰ salinity, **D:** Crystallizer pond with 300 ‰ salinity. Phenotypical differences between the certain ponds are represented by different colours of the water. The pictures **B – D** were kindly provided by Sabine Filker (University of Kaiserslautern).

2.2   RNA extraction and quality control

Total RNA from the 21 filters (7 stations x 3 replicates) was extracted using Qiagen´s AllPrep DNA/RNA Mini-Kit (Qiagen GmbH, Hilden, Germany) according to the manufacturer`s instructions (Suppl. Tab. 2) following a chemo-mechanical cell disruption by bead-beating (45 s, 30 Hz). Residual DNA was removed by DNase I (Qiagen GmbH) digestion. The concentration of the extracted and purified RNA was determined spectro-photometrically using a Nanodrop ND-1000 UV-Vis spectrometer (Nanodrop Technologies, Wilmington, DE, USA). The successful removal of all residual DNA was furthermore checked by PCR of the digested sample using primer pairs targeting the complete 18S rRNA gene and the shorter V4 region located on the 18S rRNA gene. Both primer sets were necessary because they raised together the probability to detect large parts as well as smaller fragments of residual DNA if the digestion failed in preference for larger or shorter fragments, respectively. The PCR mixture consisted of 5x GC buffer (New England Biolabs, Frankfurt am Main, Germany),

10mM of each deoxynucleotides (dNTPs, namely dATP, dGTP, dCTP, dTTP; New England Biolabs, Frankfurt am Main, Germany), nuclease-free water (Carl Roth GmbH & Co. KG, Karlsruhe, Germany), 1 U of a Phusion High Fidelity DNA polymerase (New England Biolabs, Frankfurt am Main, Germany), template RNA and either 10 µM of the primer pair EukA (5'-AACCTGGTTGATCCTGCCAGT-3'; Medlin et al., 1988; biomers.net GmbH, Ulm, Germany)/EukB (5'-TGATCCTTCTGCAGGTTCACCTAC-3'; Medlin et al., 1988; biomers.net GmbH, Ulm, Germany) or TAReuk454FWD1 (5'-CCAGCA(G/C)C(C/T)GCGGTAATTCC -3', Stoeck et al., 2010; biomers.net GmbH, Ulm, Germany)/TAReukREV3 (5'- ACTTTCGTTCTTGAT(C/T)(A/G)A -3', Stoeck et al., 2010; biomers.net GmbH, Ulm, Germany). The PCR protocol for amplifying the 18S rRNA gene consisted of an initial denaturation at 98 °C for 5 min, 30 cycles of denaturation (98 °C, 60 sec), annealing (67 °C, 60 sec) and elongation (72 °C, 120 sec) and a final elongation at 72 °C for 10 min (Tab. 2). The protocol for amplifying the V4 region is given in chapter '1.2 DNA extraction, PCR amplification and sequencing'. Both PCRs were performed on an iCycler thermal cycler (Bio-Rad Laboratories GmbH, Munich, Germany). For visualisation of possible amplified fragments and unwanted residual DNA, an agarose gel electrophoresis was done following the same procedure as mentioned above (see chapter '1.2 DNA extraction, PCR amplification and sequencing').

Table 2: PCR program for amplifying the 18S rRNA gene of the SSU rRNA with primer pair *EukA*/*EukB*

initial denaturation	98 °C	5 min	
denaturation	98 °C	60 sec	30 x
annealing	67 °C	60 sec	amplification
elongation	72 °C	120 sec	cycles
final elongation	72 °C	10 min	

The integrity and quality of the total RNA was confirmed with a RNA 6000 PicoAssay using an Agilent Bioanalyzer 2100 system (Agilent Technologies, Waldbronn, Germany) (Suppl. Fig. 2). Triplicate samples per sampling site were pooled in order to account for a potential sampling bias and also to obtain the required total RNA amount of 1.5 µg for subsequent polyA-enrichment.

2.3   Poly-A enrichment and mRNA sequencing

In order to enrich for the low abundant eukaryotic messenger RNA (mRNA) and thus eliminate prokaryotic RNA, high abundant ribosomal RNAs and other types of RNA, a polyA-selection was conducted targeting the eukaryote mRNA-specific 3'-polyadenylation during library preparation. Sequencing libraries were prepared using Illumina´s TruSeq Stranded mRNA Library Prep Kit according to the manufactures guidelines and sequenced on an Illumina NextSeq platform generating 75 bp paired-end reads by SeqIT GmbH & Co. KG (Kaiserslautern, Germany).

2.4   Bioinformatics and statistics

A systematic workflow was applied to process the read data obtained from the seven sequencing libraries. General steps included data cleaning, assembly, open reading frame (ORF) prediction, taxonomic and functional annotations, as well as statistical analyses (Fig. 6).
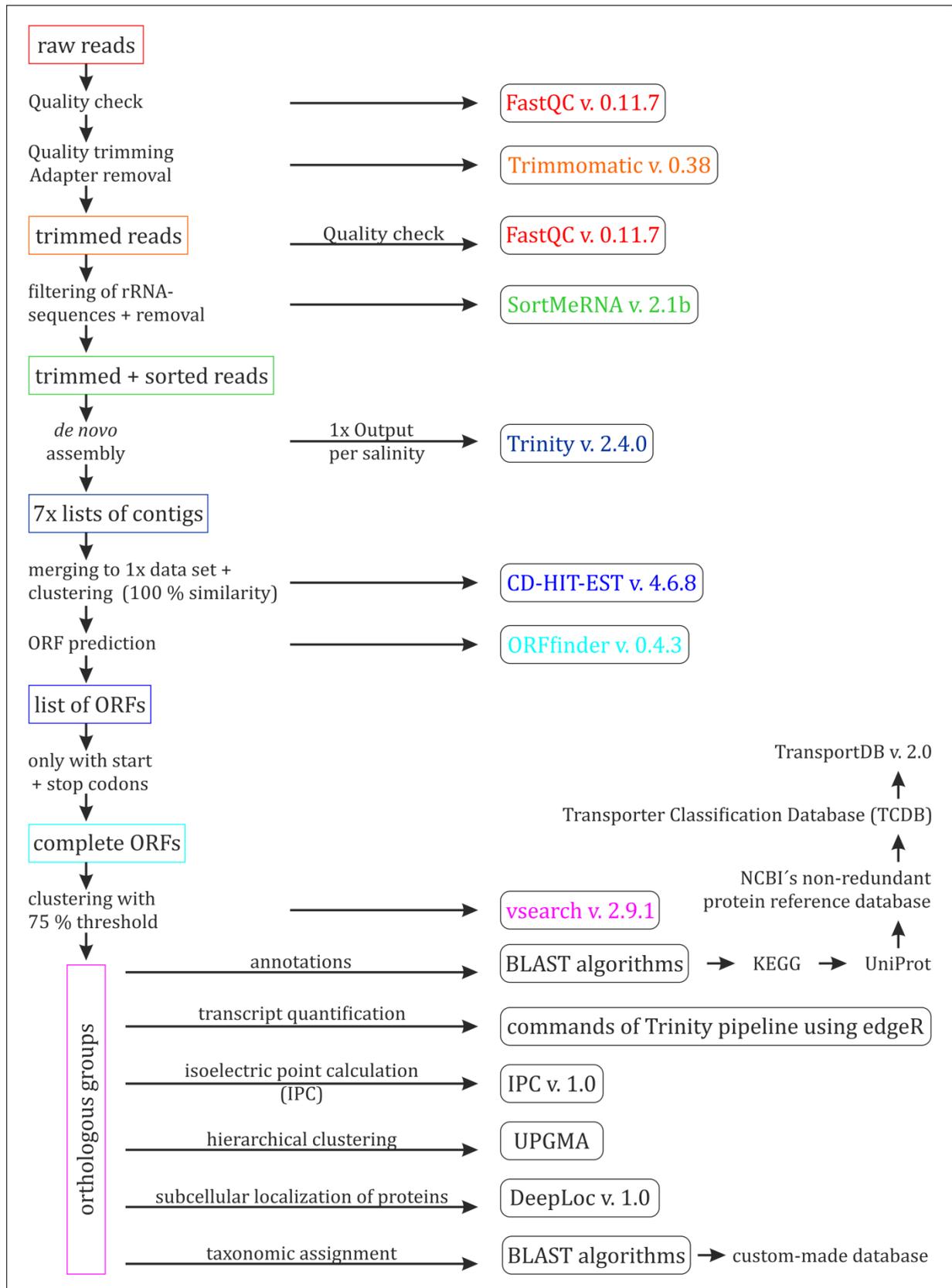
Figure 6: Overview of the bioinformatic workflow and statistical analyses of meta-transcriptomic dataset. Data rearrangements between the different software programs were done using 'R' including the packages 'dplyr', 'fossil', 'ggplot2', 'stats', 'tidyr' and 'vegan'.

2.4.1 Metatranscriptome data processing

Raw 75 bp pair-end reads were quality checked with FastQC v0.11.7 (Andrews, 2012). Then, Trimmomatic v0.38 (Bolger et al., 2014) was used to remove low quality sequences and trim adapter leftovers with the following settings: -phred33 ILLUMINACLIP:TruSeq3-PE-2.fa:2:30:10 LEADING:5 TRAILING:5 SLIDINGWINDOW:5:15 MINLEN:50. Sufficient quality filtering was confirmed again with FastQC v0.11.7 (Andrews, 2012). Reads were then screened for residual rRNA fragements using SortMeRNA v. 2.1b with default parameters (E-value = 1; Kopylova et al., 2012), which implements reference databases for prokaryotic and eukaryotic rRNA types (Fig. 6). Because of the large amount of sequence data, one *de novo* assembly per sample was generated using Trinity v2.4.0 (Grabherr et al., 2011) with default settings. After having tested and evaluated different settings of Trinity the default settings without any implemented read-normalization step appeared as most reliable.

To allow comparable transcript quantification, the seven independent assemblies were merged. Redundancy among contigs was reduced using CD-HIT-EST v. 4.6.8 (Li and Godzik, 2006; Huang et al., 2010; Fu et al., 2012) at 100 % contig similarity. Afterwards, open reading frames (ORFs) on the distinct contigs were predicted using ORFfinder v. 0.4.3 (https://www.ncbi.nlm.nih.gov/orffinder/) and translated to protein sequences with the following settings: -g 1 -s 1 -ml 100 -n true -strand both. For downstream analysis only complete ORFs were used containing a start- and stop-codon at the correct positions in the sequence. Start-codons included 'ATG', 'CTG', 'TTG' and stop-codons 'TGA', 'TAA', 'TAG'. Predicted protein sequences were clustered to produce orthologous groups, i. e. binning of possible isoforms and homologous proteins, with a 75 % threshold using vsearch v. 2.9.1 (Rognes et al., 2016) (Fig. 6).

2.4.2 Annotation

For functional annotation, representative sequences of each orthologous group (OG) were parsed through a hierarchical annotation pipeline, with unannotated OGs passing to the next annotation step (Fig. 6). This analysis included (i) mapping of OGs to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Kanehisa, 2000; Kanehisa et al., 2016a; Kanehisa et al., 2017) using GHOSTKOALA (Kanehisa et al., 2016b); (ii) sequence similarity searches consecutively against UniProt (UniProt Consortium, 2017),

NCBI´s non-redundant protein reference database (NCBI- Reference Sequence (RefSeq) database; http://www.ncbi.nlm.nih.gov/RefSeq/; O'Leary et al., 2016), the transporter classification database (TCDB; http://www.tcdb.org/; Saier et al., 2006; 2016) and TransportDB v. 2.0 (http://www.membranetransport.org/transportDB2/index.html) using either BLASTX or BLASTP (Altschul et al., 1990) with an e-value of $1e^{-5}$.

Taxonomic classifications of sequences were derived from BLASTX (Altschul et al., 1990) searches against a custom protein and cDNA reference database (Liu et al., 2018), consisting of transcriptome and genome data of more than 805 marine and hypersaline species originating from the Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP; Keeling et al., 2014). For BLASTX analyses a bit score of 90 % with an e-value of $1e^{-5}$ was applied and the five best hits per query sequence were considered. According to Hu et al. (2018), a taxonomic identity was assigned to a sequence if all extracted five top hits originated from the same phylum. If no consensus could be found, the sequence was assigned to the category 'multiple hits'.

2.4.3   Transcript quantification and differential gene expression analysis

Transcript quantification was done by following the workflow implemented in the Trinity pipeline (version 2.4.0, default parameters, Haas et al., 2013), using the short read mapper Bowtie2 (with default parameters, Langmead and Salzberg, 2012) to map the trimmed and sorted reads back to the orthologous groups. Trimmed mean of M-values (TMM) normalization (Robinson and Oshlack, 2010) was applied as the best cross sample normalization method (Dillies et al., 2013). Resulting transcript abundances were normalized per salinity to get transcript per million (TPM) values for each orthologous group according to salinity.

Differential expression analysis (DEA) on basis of the TPM values was done by comparing the transcript abundance results for each OG between the different samples using the 'R' packages 'edgeR' (Chen et al., op. 2014) and 'voom' (Law et al., 2014) with the setting --min_reps_min_cpm 2,1 --dispersion 0.1. OGs displaying 4-fold-change ($\log_2 FC \geq 2$; FDR < 0.01) in at least one sample comparison were considered as differentially expressed.

2.4.4 Partitioning of functional diversity

TPM values for each OG of each sample served as basis to investigate functional (dis)similarities between the samples. The script 'jackknifed_beta_diversity.py' implemented in QIIME v. 2.0 (Caporaso et al., 2010) was used for normalization, creation of a Bray-Curtis index based distance matrix, and subsequent unweighted pair group method with arithmetic mean (UPGMA) analysis.

In order to investigate functional changes within the samples related to osmotic response, all differentially expressed OGs were extracted, which had an annotation that could be categorized in either of one of the following metabolic functions: 'sensing', 'compatible solutes', 'compatible solute transport', 'ion transport' or 'energy'. The TPM values of OGs with the same annotated function were summed and normalized per salinity to the smallest sample size (sum of TPM per salinity) to compute a Bray-Curtis index-based distance matrix per subset. This normalization was automatically performed with the script 'jackknifed_beta_diversity.py' implemented in QIIME v. 2.0 (Caporaso et al., 2010) following the above mentioned workflow to a subsequent UPGMA analysis. The Bray-Curtis index was chosen to consider the relative abundance of a cellular function in a protistan community as the sum of TPM values of OGs sharing this function.

2.4.5 Isoelectric point calculation

To distinguish between the different proteomes of species, which can tolerate higher intracellular inorganic ion concentrations, and those, who do not have intracellular adaptations to tolerate a large amount of intracellular inorganic ions, the acidity of the cytoplasm was the dimension of choice. If the 'high-salt-in' strategy predominates, a clear shift within the community to more acidic values should be detected. Hence, the sub-cellular localization of the OGs were determined first to obtain all cytoplasmic proteins using DeepLoc-1.0, which is based on machine learning (Almagro Armenteros et al., 2017), before the acidity of the cytoplasm was calculated for each salinity (Fig. 6). The salinity specific acidity of the cytoplasm is based on the sum of the isoelectric points of each OG located in the cytoplasm using the isoelectric point calculator (IPC) v. 1.0 (Kozlowski, 2016) with default settings. Only the presence or absence of a unique OG at

certain salinity was integrated with its IPC value to calculate the isoelectric point of the whole community for this salinity.

2.4.6   Statistics

In general unless stated otherwise, all statistical analyses were conducted in R v. 3.3.2 (R Core Team, 2013) using the 'vegan' package (Oksanen et al., 2015), 'dplyr' (https://CRAN.R-project.org/package=dplyr), 'tidyr' (https://CRAN.R-project.org/package=tidyr), as well as 'ggplot2' (Wickham, 2016) for graphical representations. Statistical significant variations of the acidity (isoelectric point) of the cytoplasm of a protistan community according to the salinity were tested using a 'TukeyHSD' test implemented in the 'fossil' package (https://cran.r-project.org/package=fossil). Different letters represented mathematical significant differences between certain salinities. All R scripts are provided digital on the appended CD-Rom.

3.   Time-resolved transcriptome analysis of *Schmidingerothrix salinarum*

In this part of my PhD thesis I investigated *Schmidingerothrix salinarum´s* cellular and physiological dynamics after osmotic up-shock and during salt acclimation. The conducted time-resolved transcriptome analysis of this model ciliate was divided in an experimental setup (Fig. 7) subdivided in the chapters '3.1 Species isolation and cultivation', '3.2 Salt up-shock experiment', '3.3 Isolation of mRNA and mRNA sequencing', and sequence data analyses (Fig. 8) containing several bioinformatic analyses and statistics (see chapter '3.4 Bioinformatics analyses and statistics').

3.1   Species isolation and cultivation

The halophile, heterotroph ciliate *Schmidingerothix salinarum* (Foissner et al., 2014) was originally isolated from a solar saltern pond in Ses Salines, Mallorca, Spain (Salinas d`Es Trenc, N 39° 21' 16.11'' E 3° 0' 42.16''). Initial cultures were established by handpicking single cells in a volume of 1 µl and transferring them into sterilized salt medium (artificial seawater, ASW; Instant Ocean, Aquarium Systems, Aqua Vendi GmbH, Erkrath, Germany, for chemical composition see Atkinson and Bingman, 1997). Sterilized wheat grains were added to support the growth of indigenous food bacteria, which were isolated together with the individual *S. salinarum* cells and, thus, serve as its food source. Cultures were grown at room temperature (RT) with a 12 h light-dark cycle at 38 ‰ salinity. When cell density reached 500 cells/ml, cultures were subdivided and filled up with sterilized salt medium and a sterile wheat grain to a volume of 20 ml (Fig. 7A).
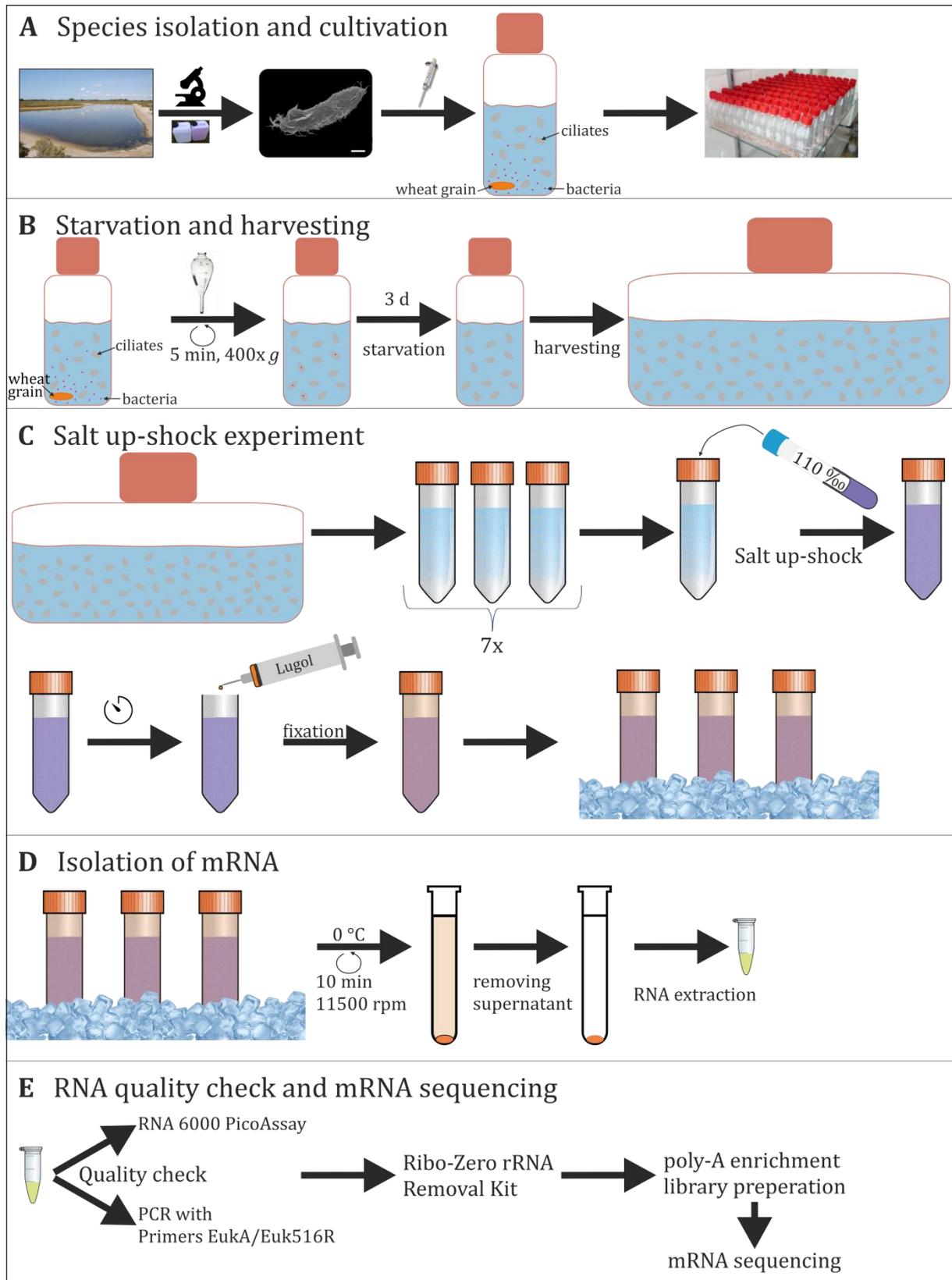
Figure 7: Overview of the lab-procedure for the ciliate *Schmidingerothrix salinarum* from species isolation to sequencing, divided in the following parts: **A**: Species isolation and cultivation; **B**: Starvation and harvesting; **C**: Salt up-shock experiment; **D**: Isolation of mRNA; **E**: RNA quality check and mRNA sequencing. In total, 21 sub-samples from 7 time points were processed after salt up-shock.

3.2   Salt up-shock experiment

To investigate the functional and metabolic response of *S. salinarum* to changes of external salinity, a salt up-shock, i. e. a sudden increase of external salinity, was applied. Prior to this, ciliate cells were concentrated by centrifugation in pear-shaped ASTM centrifuge tubes (Lenz Laborglas GmbH & Co. KG, Wertheim, Germany) for 5 min at 400x g (Roto Silenta III, Hettich, Germany), followed by removal of the bacteria-containing supernatant (Fig. 7B). Cells were checked for positive vitality under a stereo microscope (SZ-PT, Olympus, Tokyo, Japan) and then transferred into sterile salt medium. In order to reduce the number of free and ingested food bacteria, ciliate cells were starved for at least three days. After this, cells were again harvested by centrifugation as described above and transferred to sterile salt medium with a salinity of 38 ‰ (Fig. 7B). Triplicate subsamples (ca. 20,000 cells) were collected to serve as control samples for time point zero minutes (t0). The salt shock was initiated by adding a certain volume of sterile ASW medium with 110 ‰ salinity to reach a final salt concentration of 90 ‰ salinity. Triplicate subsamples (19.000-21.000 cells) were then collected after 2, 10, 30, 60, 120 and 720 minutes (Fig. 7C). To immediately stop all intracellular reactions acidic Lugols solution (10 g KI, 5 g I_2, 95 ml H_2O_{dest}, 5 ml glacial acetic acid) with a final concentration of 1 % was added to all 21 subsamples (7 time points, 3 replicates each). Fixation of intact cells was only possible using acidic Lugols solution (Suppl. Tab. 3). The cells were fixed and stored on ice (Fig. 7C).

3.3   Isolation of mRNA and mRNA sequencing

Fixed cells were centrifuged for 10 min at 11.500 rpm and 0 °C (Sorvall RC-5C Plus Superspeed Centrifuge, Thermo Fisher Scientific Inc., Germany) to get a cell pellet. The supernatant was discarded and the cell pellets were used for total RNA extraction (Fig. 7D) following the above mentioned protocol (see chapter '2.2 RNA extraction and quality control') using Qiagen´s AllPrep DNA/RNA Mini-Kit (Qiagen GmbH, Hilden, Germany) including the spectro-photometrically determination of the concentration of extracted and purified RNA as well as the verification of RNA integrity via RNA 6000 PicoAssay. Also, a PCR and agarose gel electrophoresis were done to verify the successful removal of all DNA in dependence of the above mentioned procedure (see chapter '2.2 RNA extraction and quality control') with minor modifications: Instead of primer pair EukA/EukB, the primers EukA (biomers.net GmbH, Ulm, Germany) and

Euk516R (5'-ACCAGACTTGCCCTCC-3'; Wuyts et al., 2004; biomers.net GmbH, Ulm, Germany) were used with a slightly different PCR mixture containing 120-300 ng of template RNA, 1 U of Taq-Polymerase (Axon Labortechnik, Kaiserslautern, Germany), 10x CoralLoad PCR buffer (Axon Labortechnik, Kaiserslautern, Germany), 200 µmol/l of each dNTPs (dATP, dGTP, dCTP, dTTP; New England Biolabs, Frankfurt am Main, Germany) and 0.25 µmol/l of each oligonucleotide primer. The final volume was adjusted to 50 µl with nuclease-free water (Carl Roth GmbH & Co. KG, Karlsruhe, Germany). Due to a variation of the used primers, the PCR protocol for amplifying the 18S rRNA gene from the above mentioned chapter ('2.2 RNA extraction and quality control') was used with minor modifications: The denaturation temperature was 95 °C and the annealing occurred at 57 °C (Tab. 3).

Table 3: PCR program for amplifying the 18S rRNA gene of the SSU rRNA with primer pair *EukA/Euk516R*

initial denaturation	95 °C	120 sec	
denaturation	95 °C	30 sec	30 x
annealing	57 °C	30 sec	amplification
elongation	72 °C	60 sec	cycles
final elongation	72 °C	5 min	

Total RNA samples were treated with Illuminas Ribo-Zero rRNA Removal Kit for Bacteria to reduce the amount of RNA transcripts potentially deriving from food bacteria. Afterwards, cytosolic mRNA from the ciliate was enriched by poly-A selection. Sequence library preparation was done using Illuminas TruSeq Stranded mRNA Library Prep Kit (Fig. 7E) following the manufactures guidelines. Sequencing of the 21 libraries occurred on an Illumina NextSeq platform by SeqIT GmbH & Co. KG (Kaiserslautern, Germany), producing 75 bp paired-end reads.

3.4   Bioinformatic analyses and statistics

In dependence of the systematic workflow for the metatranscriptome data (see chapter '2.4 Bioinformatics and statistics'), a modified systematic workflow was applied to process the read data obtained from the 21 sequencing libraries from the transcriptome of *S. salinarum*. General steps included data cleaning, assembly, open

reading frame (ORF) prediction, differential expression analysis (DEA), functional annotations and statistical analyses (Fig. 8).
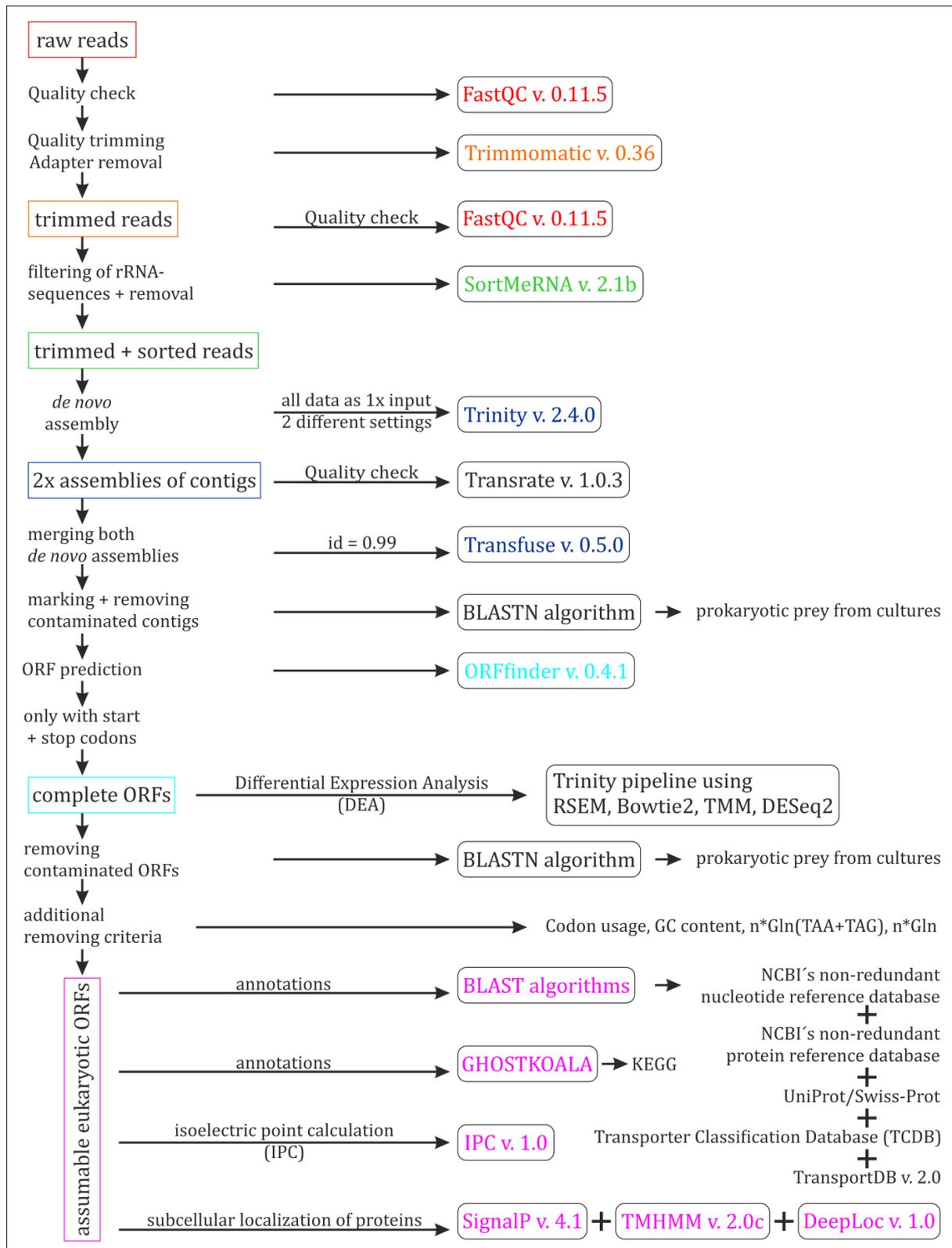
Figure 8: Overview of the bioinformatic sequence data analyses and statistics of all 21 sub-samples from the mRNA sequencing of *S. salinarum*. Data rearrangements between the different software programs were done using 'R' including the packages 'dplyr', 'fossil', 'ggplot2', 'stats', 'tidyr' and 'vegan'.

3.4.1 Time-dependent transcriptome data processing

The raw 75 bp paired-end reads were processed as mentioned above (see chapter '2.4.1 Metatranscriptome data processing') including quality check, adapter trimming and removal of rRNA fragments. High-quality reads of each library were pooled and subjected to a *de novo* assembly using Trinity v. 2.4.0 (Grabherr et al., 2011) with two different settings, i. e. with and without read normalization (Fig. 8; exact commands are listed in the digital appendix). The quality of both assemblies was compared using Transrate v. 1.0.3 (Smith-Unna et al., 2016). Both assemblies were then merged together using Transfuse v. 0.5.0 (https://github.com/cboursnell/transfuse) with id=0.99 in order to increase assembly quality (Fig. 8).

To remove any remaining prokaryotic contaminations that overcame the molecular processing before, obtained contigs were aligned to prokaryotic genomes originating from bacterial prey in the culture medium using BLASTN (Altschul et al., 1990). Contigs with any significant match to non-eukaryotes were marked in the dataset and removed after differential expression analysis (DEA). In the next step, ORF prediction and selection of complete ORFs was done as described above (see chapter '2.4.1 Metatranscriptome data processing') with the exception of a ciliate specific codon usage, 'ATG' as exclusively start- and only 'TGA' as stop-codon (Cavalcanti et al., 2004).

3.4.2 RNAseq expression analysis

To prevent mismatches during abundance estimation, transcript quantification was done before potential prokaryotic contaminations were removed from the final dataset because trimmed and sorted reads from potential bacterial contaminations would not have been mapped correctly if unwanted contigs or ORFs were discarded to early from the dataset. The Trinity pipeline (version 2.4.0, default parameters, Haas et al., 2013) including the abundance estimation via RSEM (Li and Dewey, 2011) was used to quantify transcript abundances, and the normalization procedure TMM was applied for cross-sample normalization (see chapter '2.4.3 Transcript quantification'). In order to detect real biological differences and exclude possible batch effects from the sequencing machinery, an additional Trinity-integrated perl script (/software/trinityrnaseq/2.4.0/Analysis/DifferentialExpression/remove_batch_effects_ from_count_matrix.pl) was performed before the final steps of differential expression

analysis (DEA) finished. During the DEA, all time points were compared to the control group (t0) using DESeq2 (Love et al., 2014). ORFs having a minimum counts-per-million (cpm) value of 0.1 in at least two subsamples, a FDR < 0.001 and a 4-fold-change or higher ($\log_2 FC \geq 2$) compared to the control group (t0) were considered differentially expressed.

3.4.3 Removal of prokaryotic contaminations

To remove prokaryotic contaminations that may have passed the applied quality filter, several steps and methods were implemented. First, the ORFs of (above mentioned) marked contigs with any significant match to non-eukaryotes, possibly originating from bacterial prey as food source in the culture medium, were removed from the final dataset. Second, all complete ORFs were similarly mapped to known bacterial genomes originating from bacterial prey using BLASTN (Altschul et al., 1990) and together with all incomplete ORFs discarded, if necessary. Third, all remaining ORFs were analysed according to the following criteria: general codon usage, GC content, number of glutamines (Gln) coded by TAA or TAG (Martindale, 1989; Heaphy et al., 2016), and total number of glutamines (Gln) to distinguish between eukaryotic sequences of the investigated organism *Schmidingerothrix salinarum* and other unwanted contaminations (Fig. 8).

The codon usage of each ORF was analysed by a custom script (exact commands are listed in the digital appendix) via R v. 3.3.2 (R Core Team, 2013) and compared to the information shared in the codon usage database (http://www.kazusa.or.jp/codon/). The GC content was calculated using a perl script (http://alrlab.research.pdx.edu/aquificales/scripts/get_gc_content.pl). The numbers of glutamines were counted according to their translated triplet. Additionally, the nucleotide sequences of the ORFs and their corresponding protein sequences were compared to GenBank (NCBI-GenBank Flat File Release 222.0; Benson et al., 2005) and NCBI´s non-redundant nucleotide reference database (NCBI- Reference Sequence (RefSeq) database; http://www.ncbi.nlm.nih.gov/RefSeq/; O'Leary et al., 2016) as well as NCBI´s non-redundant protein reference database (NCBI- Reference Sequence (RefSeq) database; http://www.ncbi.nlm.nih.gov/RefSeq/; O'Leary et al., 2016) using BLASTN or BLASTP (Altschul et al., 1990) searches based on an e-value of 1e^{-5} to

determine whether it is more likely that the sequence belongs phylogenetically to a eukaryote or not. On the basis of these criteria, ORFs with assumable eukaryotic origin were kept, whereas potential prokaryotic ORFs were discarded from the final dataset (Fig. 8).

To verify the decision based on the above mentioned criteria and for defining correct thresholds for each individual criterion, ORFs were additionally aligned via BLASTN (Altschul et al., 1990) against complete genome sequences of *Schmidingerothrix salinarum* (unpublished draft-genome data; provided in the digital appendix upon request). Matching sequences were thus collected to extract GC content values and codon usage parameters.

3.4.4 Functional annotation

ORFs with assumable eukaryotic origin were annotated according to the above mentioned approach (see chapter '2.4.2 Annotation') to the following reference databases (with decreasing importance, see Fig. 8): UniProt/Swiss-Prot (UniProt Consortium, 2017), NCBI´s non-redundant protein (nr) reference database (NCBI-Reference Sequence (RefSeq) database; http://www.ncbi.nlm.nih.gov/RefSeq/; O'Leary et al., 2016), transporter classification database (TCDB; http://www.tcdb.org/; Saier et al., 2006; Saier et al., 2016), TransportDB v. 2.0 (http://www.membranetransport.org/transportDB2/index.html). Additionally, GHOSTKOALA (Kanehisa et al., 2016b) was used to map the amino acid sequences of all ORFs to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Kanehisa, 2000; Kanehisa et al., 2016a; Kanehisa et al., 2017).

Differentially expressed ORFs of *S. salinarum* with at least onetime 4-fold-change ($log_2FC \geq 2$) compared to the control (t0) were categorized according to their functional annotation in order to detect osmotically relevant intracellular dynamics on a time-dependent scale. This functional partitioning of differentially expressed ORFs into the categories 'sensing', 'compatible solutes', 'compatible solute transport' and 'ion transport' permits the integration of ORFs with identical functions by calculating their expression compared to control (t0) as the mean of the single protein values.

3.4.5   Isoelectric point calculation

To investigate the whole proteome of *Schmidingerothrix salinarum* and to detect potential haloadaptation mechanisms, isoelectric points of all complete eukaryotic ORFs were calculated using the isoelectric point calculator (IPC) v. 1.0 (Kozlowski, 2016) with default settings. To distinguish between cytoplasmic proteins and other cell compartments, predicted proteins were first analysed with SignalP v. 4.1 (Petersen et al., 2011) for discriminating between signal peptides and transmembrane regions. Then, the orientation of a predicted transmembrane region was checked with TMHMM v. 2.0c (Sonnhammer et al., 1998; Krogh et al., 2001) and the final prediction of the subcellular localization of the protein was executed using DeepLoc-1.0 (Fig. 8), which contains a neuronal network algorithm based on machine learning (Almagro Armenteros et al., 2017). The acidity of the different cell compartments of *S. salinarum* was calculated as frequency distribution of all proteins located in that cell compartment according to their isoelectric point. If the majority of proteins is shifted to more acidic values, this cell compartment is likely more acidic and it is assumable that more intracellular inorganic ions can be tolerated. To evaluate the findings for the acidity of *S. salinarum*, all proteins in the NCBI´s non-redundant protein reference database (NCBI- Reference Sequence (RefSeq) database; http://www.ncbi.nlm.nih.gov/RefSeq/; O'Leary et al., 2016) of the known species *Salinibacter ruber*, *Hortaea wernickii* and *Oxytricha trifallax* were extracted and their isoelectric point was calculated as it is mentioned above for *S. salinarum*. The frequency distributions of the proteins of the four species according to their isoelectric point were compared as indicator for the acidity of their proteome.

3.4.6   Statistics

In general unless stated otherwise, all statistical analyses were conducted as described above (see chapter '2.4.6 Statistics'). Additionally, graphical representations were created by a 'PtR'-script of the Trinity pipeline (/software/trinityrnaseq/2.4.0/Analysis/DifferentialExpression/PtR). All R scripts are provided digital on the appended CD-Rom.

III.   Results

1. eDNA metabarcoding of protistan plankton communities in the Baltic Sea

Environmental protistan DNA was collected in surface water layers along a 600 km transect in the Baltic Sea and the obtained protistan eDNA metabarcodes were analysed via high-throughput next-generation sequencing to provide reliable data for protistan plankton thriving in low-salt habitats and subsequently to reveal their diversity, community composition and distribution patterns.

1.1 Environmental parameters and sequence data overview

The surface water of the sampled transect in the Baltic Sea was characterized by a pronounced salinity gradient from west to east (Fig. 9 and Suppl. Tab. 1). In the Kiel Bight, salinity ranged from 17.3 to 22.1 ‰, with its maximum at station X4. More eastwards, salinity decreased significantly within Mecklenburg Bay and Arkona Basin from 19.4 down to 8.0 ‰ (Fig. 9). In the Bornholm Basin and Gotland Basin, the salinity fluctuated between 7.9 (station B6, Bornholm Basin) and 8.4 ‰ (station X17, Gotland Basin). Due to this salinity gradient, the sampling transect could be divided into three salinity zones (Fig. 9): First, the Western Baltic Sea including Kiel Bight and Mecklenburg Bay with the highest salinities ranging from 17.3 to 22.1 ‰ covering the stations X1 to X5. Second, the western Arkona Basin with its steep decline of salinity from 16.8 down to 8.5 ‰ within the stations X6 to X11. Third, the eastern Arkona Basin together with Bornholm Basin and eastern Gotland Basin including stations A6 to X17 with lower and little fluctuating salinities around 7.9 to 8.4 ‰.

Measurements of temperatures and oxygen concentrations showed no distinctive gradient along the sampling transect (Fig. 9). Temperatures were fluctuating between 4.6 °C at station X16 (Gotland Basin) and 7.4 °C at station X3 (Kiel Bight) (Suppl. Tab. 1) resulting in a slight decrease from west to east (Fig. 9). Oxygen concentrations remained rather stable along the transect ranging from 7.4 mg/l (station V5, Kiel Bight) to 9.4 mg/l (station X11, Arkona Basin). In general, higher oxygen concentrations were found at lower salinities and the measured surface water temperature was lowest in the eastern sample stations of the Baltic Proper (Fig. 9 and Suppl. Tab. 1).
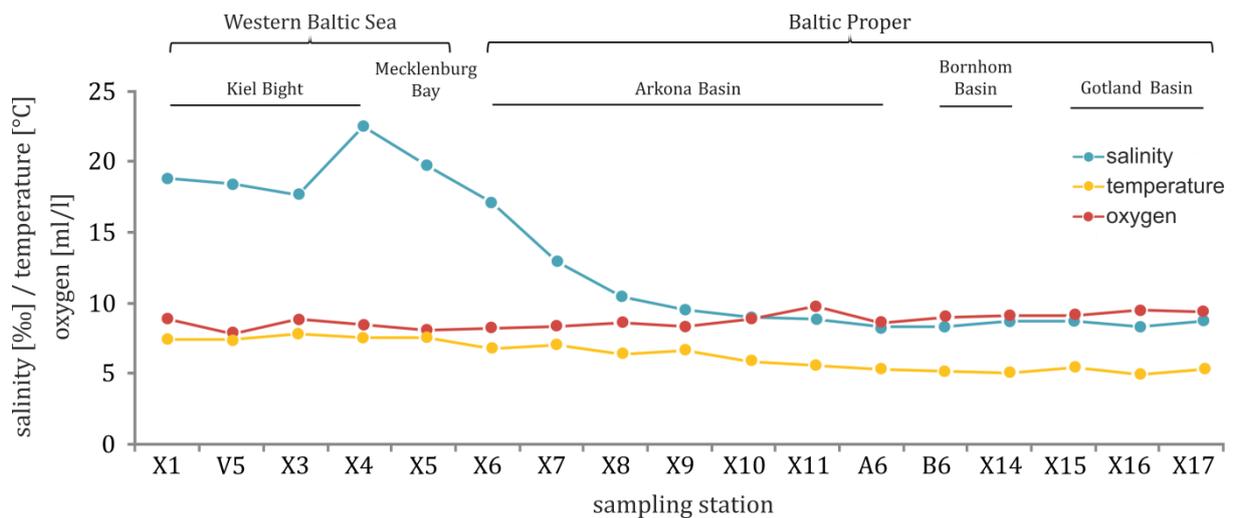
Figure 9: Salinity, temperature and oxygen profiles in surface water layers along the sampling transect showed a pronounced salinity gradient from west to east, a slight temperature decrease from west to east and almost rather stable oxygen concentrations with little higher values at lower salinities. Raw data can be found in Suppl. Tab. 1. The profiles were slightly modified from Filker et al. (2019).

Sequencing yielded on average 164,821 reads per sample (min: 88,374 station X1, Kiel Bight; max: 260,175 station X4, Kiel Bight). Quality control and elimination of potential chimeras, singletons/doubletons and non-protistan sequences reduced the read number to on average of 139,041 reads per sample, with its maximum at station X4 (227,612) and its minimum at station X1 (81,007), both samples being located in Kiel Bight (Suppl. Tab. 1). Sequence reads were grouped into an average of 343 OTUs per sample (min: 65 OTUs station X1, Kiel Bight; max: 552 OTUs station X9, Arkona Basin) (Suppl. Tab. 1). Rarefaction curves of all samples approached the asymptote indicating sample saturation (Suppl. Fig. 3).

1.2 Alpha-diversity patterns

Investigating alpha-diversity patterns, the Simpson index for whole protistan communities and also the effective number of species (exp(H')) increased significantly with decreasing salinity from 22.1 ‰ to 7.9 ‰ (Fig. 10). Pearson correlations confirmed, with r = 0.9 (Simpson) and r = 0.83 (exp(H')), the linear regression models (p < 0.01 in both cases) (Tab. 4). In contrast, rarefied OTU richness varied independently of salinity (Fig. 10).

Similar to salinity, temperature correlated with the effective number of species and the Simpson index, but not with rarefied OTU richness. Pearson correlations

provided significant r = 0.75 for exp(H') and r = 0.84 for Simpson with p < 0.01 (Tab. 4). In contrast, oxygen had a significant influence on the effective number of species and Simpson index, but Pearson r was too low in both correlations to be informative (Tab. 4).
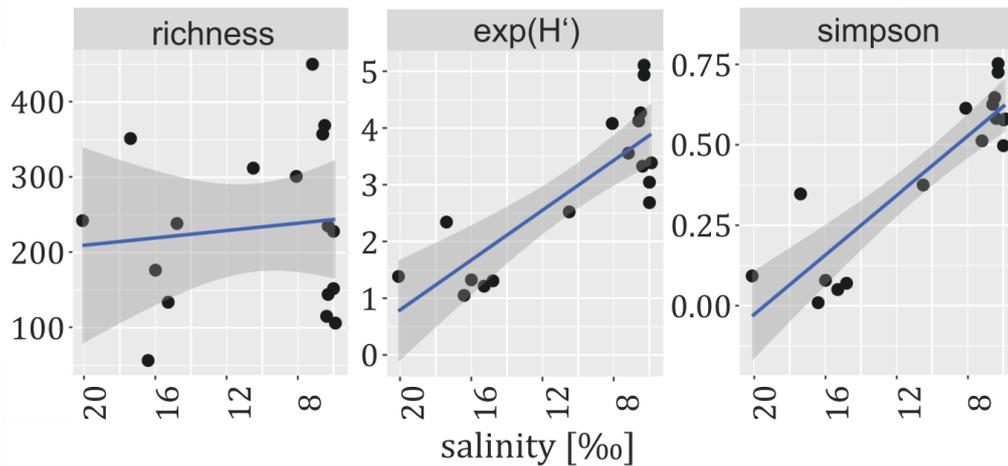
Figure 10: Measurements of protistan community alpha-diversity in terms of rarefied OTU richness, effective number of species and Simpson index as a function of salinity in surface water samples collected along the west–east transect of the Baltic Sea. Blue line = linear regression fit; shaded area = 95 % confidence interval. Additional Pearson correlations were calculated (Tab. 4). The figure was slightly modified from Filker et al. (2019).

Table 4: Pearson correlations of OTU richness and OTU diversity (exp(H') and Simpson) with environmental parameters of surface samples. Correlations are considered significant, if p < 0.05 and r > 0.75. The table was was slightly modified from Filker et al. (2019).

		richness		exp(H')		simpson	
		p	r	p	r	p	r
	salinity	0.6691	-0.1119	< 0.01	-0.8281	< 0.01	-0.8958
surface	temperature	0.6775	0.1089	< 0.01	-0.7548	< 0.01	-0.8428
	O_2	0.567	-0.1494	0.0208	0.5548	0.0123	0.5915

1.3   Protistan community composition and beta-diversity patterns

Along the west-east transect, Baltic Sea surface waters were characterised by a pronounced gradient in protistan community composition (Fig. 11). The most significant change occurred between the sampling stations X7 and X8 in the western Arkona Basin, where salinity dropped from >12 ‰ to <10 ‰. In the western Baltic stations X1 – X7 containing Kiel Bight, Mecklenburg Bay and the western Arkona Basin, dictyochophyceae (= silicoflagellates) accounted markedly for the highest proportion of OTUs (mean: 67.8 % ± 9.6). In contrast, dictyochophyceae contributed on average only for 2 % (± 1.1) in the eastern Arkona Basin, the Bornholm and the Gotland Basin at all

stations east of X7 (Fig. 11). With an opposite trend, the proportion of dinophyceae exceeded all other taxon groups by far in the Baltic Proper stations X8 – X17 (mean: 87.3 % ± 2.6) and counted for < 30 % in the western stations X1 – X7 (Fig. 11). Bacillariophyceae (diatoms) and Chlorophyta OTUs had a higher relative proportion in the Baltic Sea stations X1–X7 than in all others strengthening the difference between sampling stations X7 and X8. Interestingly, the transition for the relative contribution of ciliate OTUs was not between stations X7 and X8 as for the aforementioned taxonomic lineages with predominantly pigmented species. Instead, a notable change in the relative contribution of this heterotrophic taxon group occurred in the Arkona Basin between stations X10 and X11. The average relative contribution of ciliate OTUs from stations X11 – X17 was more than twice as high (mean: 5.4 % ± 2.4) as from the stations X1 – X10 (2.4 % ± 0.7). Apart from the most dominant taxon groups, dinophyceae and dictyochophyceae, all other taxon groups accounted only for < 5 % to all OTUs in each sample (Fig. 11).
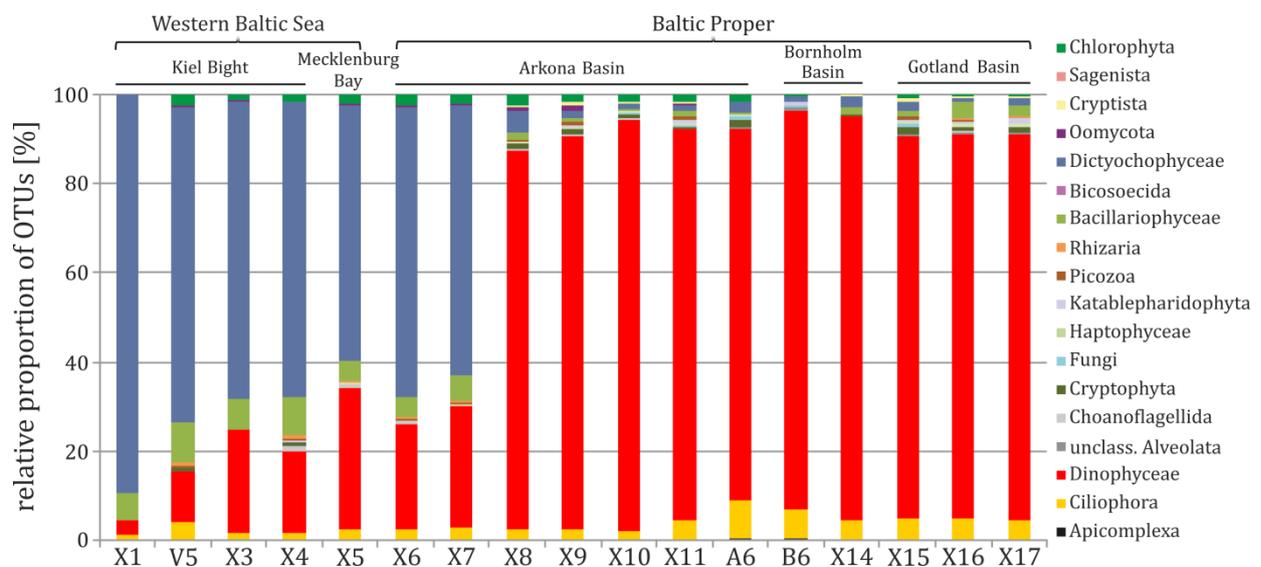
Figure 11: Relative proportion of OTUs assigned to higher taxonomic ranks collected along the west–east transect of the Baltic Sea in surface water samples. Sample names are given below representing significant differences between sample stations X7 and X8, both located in the area near Darß. The figure was slightly modified from Filker et al. (2019).

To statistically compare the above observed differences between samples, beta-diversity analysis was based on individual OTUs to provide a higher resolution for each sampling site. NMDS analyses identified three large clusters (Fig. 12). The first cluster included the western Baltic stations X1 – X7 covering salinities from 12.5 to 22.1 ‰. The second cluster contained stations X8 – X11 with lower salinities ranging from 8.5 –

10.1 ‰, whereas the third cluster consisted of the eastern stations X14 – X17 together with stations A6 (Arkona Basin) and B6 (Bornholm Basin) covering salinities between 7.9 ‰ and 8.4 ‰. To identify relevant factors for this grouping, environmental variables were fitted into the NMDS analyses (Fig. 12). Salinity was identified as the strongest selection factor to explain the separation of stations X1 – X7 (cluster 1) from the other two clusters (stations X8 – 17). The separation of cluster 2 with the stations X8 – X11 from the stations X14 – X17, A6 and B6 of cluster 3 was explained with temperature and oxygen as the stronger explanatory variables compared to salinity (Tab. 5).
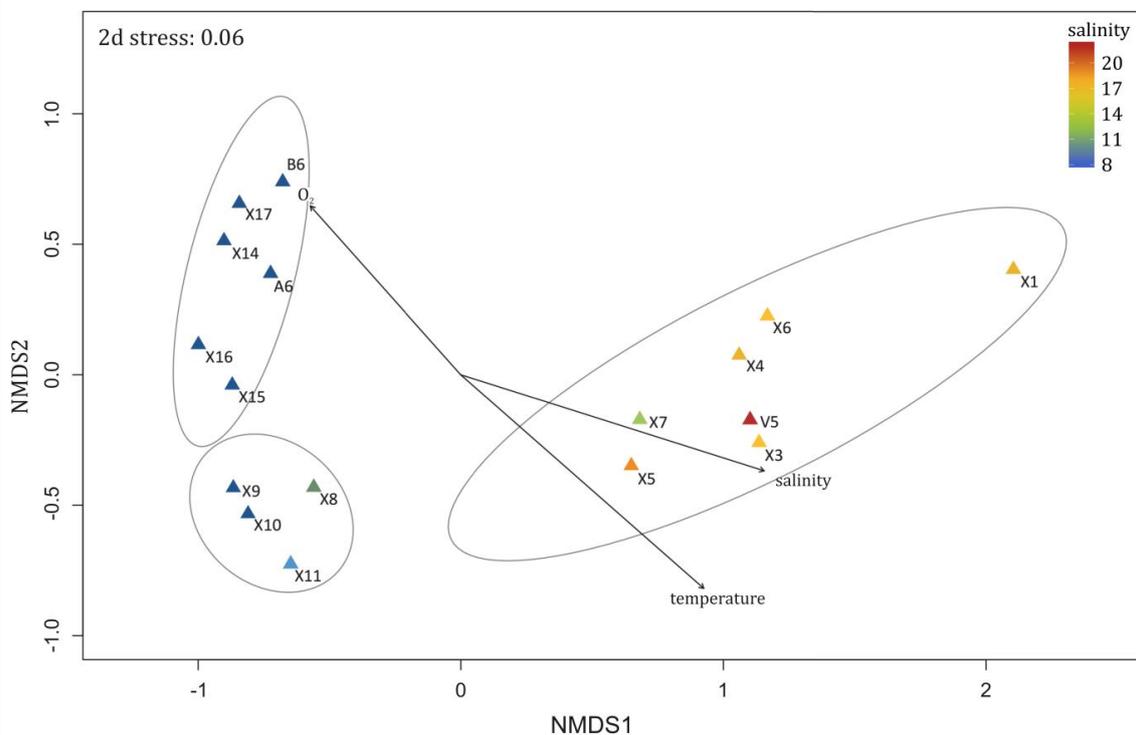
Figure 12: NMDS plot of protistan plankton community similarities collected along a salinity gradient on the water surface in the Baltic Sea. Vectors are significant explanatory environmental gradients relating community structures. Samples are colour-coded according to the measured salinity at each sampling station. Communities cluster according to salinity. Also, temperature and oxygen are relating environmental variables explaining the community clusters at lower salinities. The plot was slightly modified from Filker et al. (2019).

Table 5: Environmental parameters fitted into nonmetric multidimensional scaling analyses (NMDS) of surface water samples (Fig. 12)). P-values (Pr(>r)) < 0.05 are considered as significant effect. The table was slightly modified from Filker et al. (2019).

	NMDS1	NMDS2	r^2	Pr(>r)	
salinity	0.953	-0.304	0.857	0.001	***
temperature	0.749	-0.663	0.889	0.001	***
oxygen	-0.663	0.748	0.435	0.021	*

As revealed by a one-way ANOVA, alpha-diversity measurements, i. e. OTU richness, Simpson index and effective number of species, differed significantly ($p < 0.05$) between the three clusters (Suppl. Fig. 4B).

To verify that indeed salinity itself is the main structuring factor for protistan communities and that geographic distance had no considerable influence on the community composition, the dissimilarity between protistan communities was tested according to salinity and geographic distance separately. The dissimilarity between protistan communities from two distinct sampling sites increased significantly with increasing differences in salinity between the same two sampling sites ($p < 0.01$; $r = 0.7$; Fig. 13).
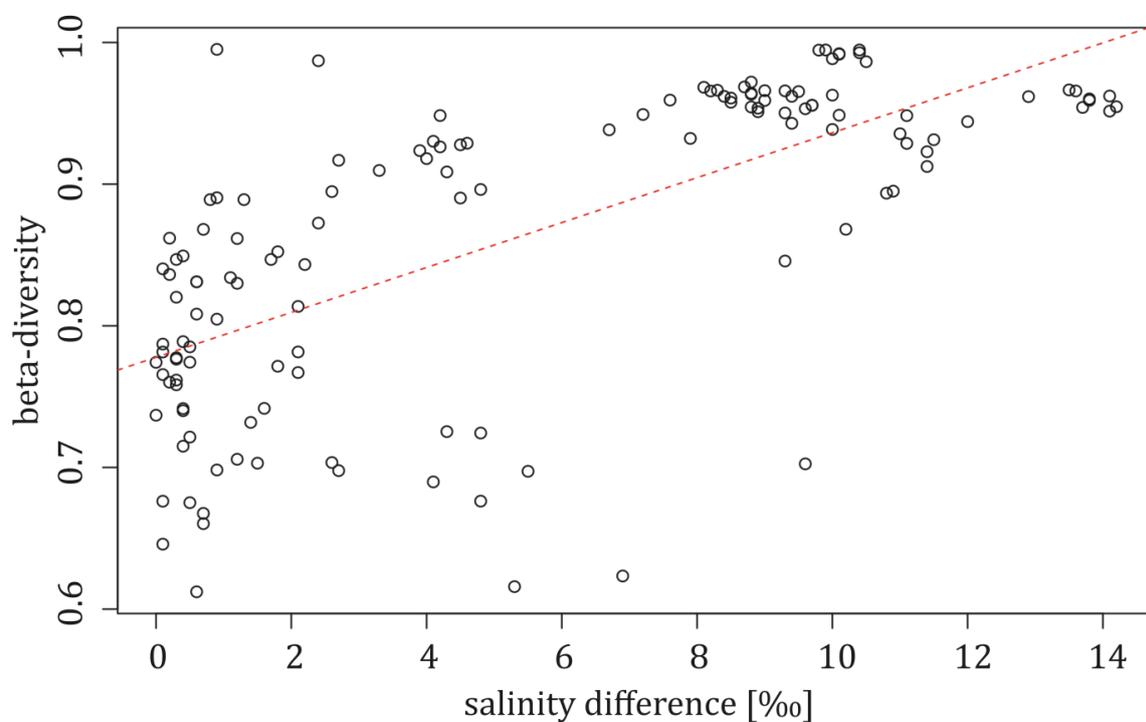
Figure 13: Regression plot from Pearson correlation of Bray–Curtis distances as a measure of beta-diversity and changes in salinity. A significant positive correlation was observed between the dissimilarity of protistan community structures and changes in salinity. If the salinity difference between a sample pair increases, community dissimilarity of the same sample pair co-increased. The plot was slightly modified from Filker et al. (2019).

A partial Mantel test was performed to assess the importance of geographic distance because salinity increased linearly along the West-East sampling transect. While the significance of the correlation between geographic distance and community dissimilarity decreased from $r = 0.5$ ($p < 0.01$) to $r = 0.2$ ($p < 0.05$) when controlled for differences in salinity (Suppl. Fig. 5), the correlation between community-dissimilarity

and salinity differences increased further after controlling for geographic effects (p < 0.01, r = 0.8). Therefore, salinity itself could be revealed as the major driving force for protistan community structures in surface water layers along the west–east sampling transect in the Baltic Sea rather than geographic distance.

1.4 Salinity-dependent transition boundaries for protistan species turnover

To control whether observed clustering patterns of the Baltic Sea protistan communities are robust also on a large biogeographic distance scale and to compare them with communities from higher salinities, the 17 Baltic Sea datasets were merged with 23 datasets of the study conducted by Filker et al. (2017). These datasets comprised sequence data of protistan communities thriving between 39 and 440 ‰ salinities, originating from solar salterns located in France, Portugal, Spain, Chile, Argentina and on Cape Verde islands (Suppl. Tab. 4). Combining these datasets extended the investigated salinity gradient from almost freshwater (8 ‰ salinity) to extreme hypersaline conditions (440 ‰ salinity). OTU richness, i. e. number of OTUs in a sample, was used as estimation for alpha-diversity (Fig. 14B). For beta-diversity analyses, the incidence-based Jaccard index was calculated and Jaccard similarity values were transformed into a distance matrix. The resulting distance matrix was used as basis for hierarchical clustering analysis in terms of UPGMA (Fig. 14A).

The integration of both datasets provided six distinct clusters that could be linked to specific salinity categories. The categories ranged from 8 – 10 ‰ (cluster 1), 12 – 22 ‰ (cluster 2), 39 – 90 ‰ (cluster 3), 120 – 240 ‰ (cluster 4), 270 – 320 ‰ (cluster 5) and 330 – 440 ‰ (cluster 6; Fig. 14). The alpha-diversity (taxonomic richness) of the protistan plankton communities within the respective clusters increased from the overall lowest diversity found in the cluster 8 – 10 ‰ to the highest diversity in cluster 120 – 240 ‰. In salinities of 270 ‰ and higher, the taxonomic richness decreased again. The taxonomic richness is, however, still higher compared to the taxonomic richness of protistan communities in the brackish water zone (clusters 8 – 10 ‰ and 12 – 22 ‰; Fig. 14B). Complementary to former findings of three transition boundaries for protistan plankton in hypersaline environments (Filker et al., 2017), protistan community shifts were observed between 90 and 120 ‰, 240 and 270 ‰, 320 and 330 ‰ for hypersaline waters. Additional community shifts for protistan plankton

existed similarly between brackish (8 – 10 ‰) and brackish – marine (12 – 22 ‰) environments as well as between 22 ‰ and 39 ‰ salinity (Fig. 14A). Considering the bootstrap values and the place of their nodes (Fig. 14), the most distinct and well supported community shifts occurred between 10 and 12 ‰ (transition from brackish to brackish – marine), 22 and 39 ‰ (transition from brackish – marine to marine low-hypersaline), 90 and 120 ‰ (transition from marine low-hypersaline to mid-hypersaline) and between 240 and 270 ‰ salinity (transition from mid-hypersaline to extreme-hypersaline).
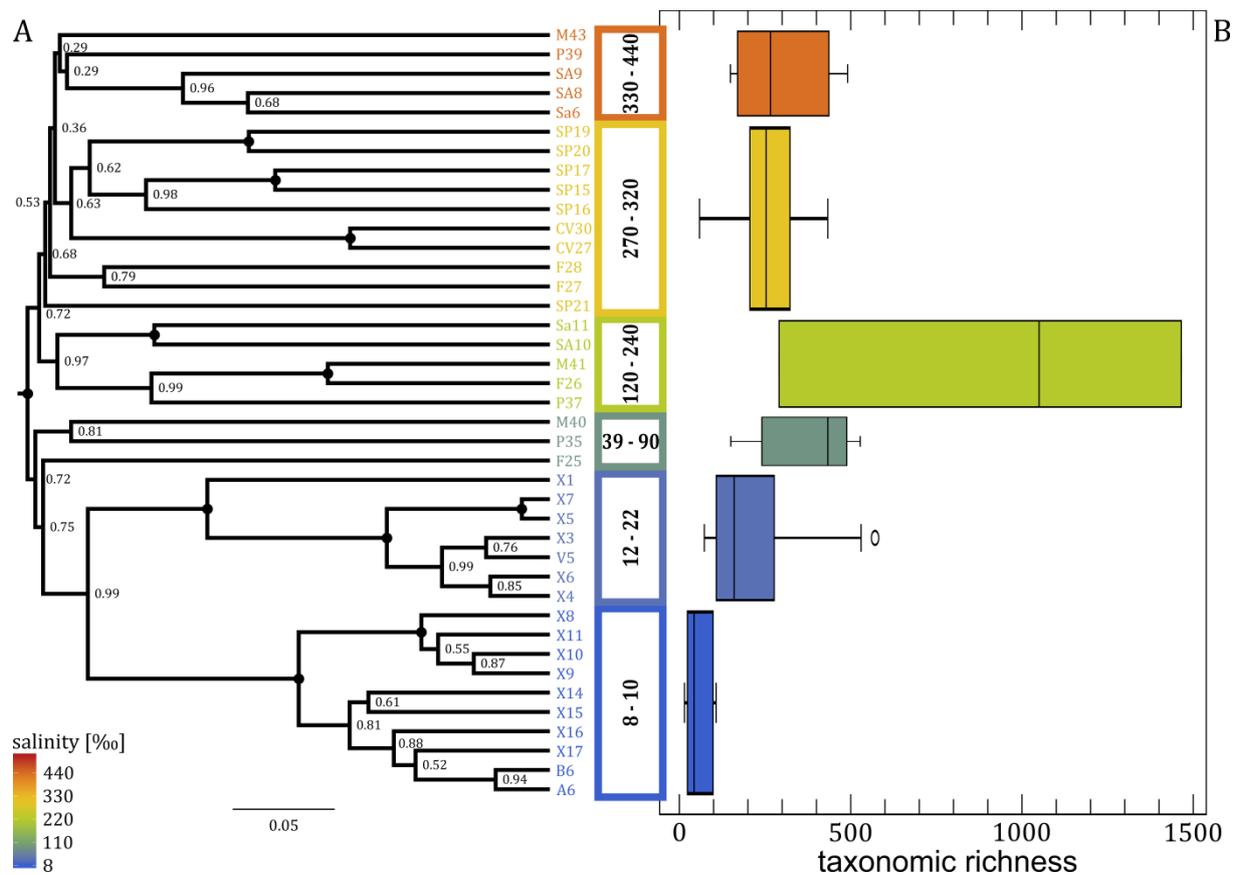
Figure 14: Protistan plankton community composition is structured in six different groups according to their salinity. **A**: UPGMA clustering of incidence-based Jaccard similarity. Bootstrap values indicate the strength of relationships in the dendrogram (• = 1). **B**: One-way ANOVA of the taxonomic richness between salinity classes (alpha = 0.05, $p < 0.01$).

In order to explain the reasons for the observed shifts in protistan plankton communities along salinity gradients, the following chapter focuses on the investigation of metabolic properties associated to salt-stress response of protistan communities living at salinities ranging between 40 ‰ to 380 ‰.

2. Analyses of microeukaryotic gene expression patterns along a salinity gradient

To test whether the observed shifts in protistan community composition of subject 'I' can be linked to the usage of different haloadaptation strategies, microeukaryotic gene expression patterns were investigated by conducting environmental, comparative meta-transcriptome analyses of natural protistan communities thriving between 40 and 380 ‰ salinity. This approach allowed not the assignment of a specific strategy to a particular organism, but enabled on the basis of expression differences to analyse potential physiological functions of whole protistan communities in response to salinity without cultivation-biases.

2.1 Sequence data overview and functional annotation

Sequencing of the seven mRNA libraries (one library per sampled pond) yielded between 126,972,520 and 169,809,332 reads per sample. Quality cleaning including adapter trimming, quality control and removal of ribosomal transcripts reduced read numbers to 114,688,388 – 159,892,094 reads per sample (Tab. 6). These reads were *de novo* assembled into 193,142 to 936,875 contigs per assembly, and encoded between 518,428 and 1,503,558 predicted proteins (Tab. 6).

Table 6: Sequence data overview according to the different salinities of the sampled ponds as 'sample ID' including the number of contigs per assembly, the number of predicted proteins per salinity and the average number of predicted proteins per contig.

Salinity of the sampled pond	Number of raw reads	Number of trimmed reads	Number of trimmed and sorted reads	Number of contigs per *de novo* assembly	Number of predicted proteins per assembly	Average number of predicted proteins per contig
40 ‰	144,951,548	139,622,902	136,710,820	936,875	1,370,303	1.5
110 ‰	169,809,332	162,596,936	159,892,094	719,816	1,503,558	2.1
150 ‰	145,132,952	140,542,804	135,638,754	487,975	1,116,174	2.3
200 ‰	129,759,828	124,978,216	122,609,394	469,838	1,136,168	2.4
240 ‰	161,567,822	155,382,776	153,189,082	193,142	518,428	2.7
300 ‰	126,972,520	122,526,700	114,688,388	492,130	1,071,834	2.2
380 ‰	154,005,802	148,204,452	128,592,418	262,256	762,013	2.9

The average number of predicted proteins per contig varied between the samples and more than one protein per contig can be found due to the consideration of different codon usages by different organisms, technical shortcomings and biases caused by creating *de novo* assemblies. For these reasons, the predicted proteins were clustered to orthologous groups (OGs) with a 75 % threshold (see chapter '2.4.1 Metatranscriptome data processing') to merge similar or identical proteins. The clustering resulted in 1,165,559 OGs, of which 338,510 could be annotated with KEGG (79,349), UniProt (234,754), NCBI´s non-redundant protein reference database (18,587), TCDB (4,808) and TransportDB v. 2.0 (1,012). The amount of functionally annotated OGs differed according to different salinity samples (Fig. 15). On average, only 28 % (327,164) of all OGs could be mapped to a database entry with a functional reference sequence. The most annotated OGs (31 %) were found for the 40 ‰ salinity sample, whereas the most unannotated OGs were observed for the 200 ‰ and 240 ‰ salinity samples (83 % and 82 %). OGs shared between the different salinity samples could be functionally annotated to 30 % (Fig. 15). In general, fewer OGs from hypersaline samples could be annotated to any functional protein reference compared to the marine sample (40 ‰ salinity sample; Fig. 15).
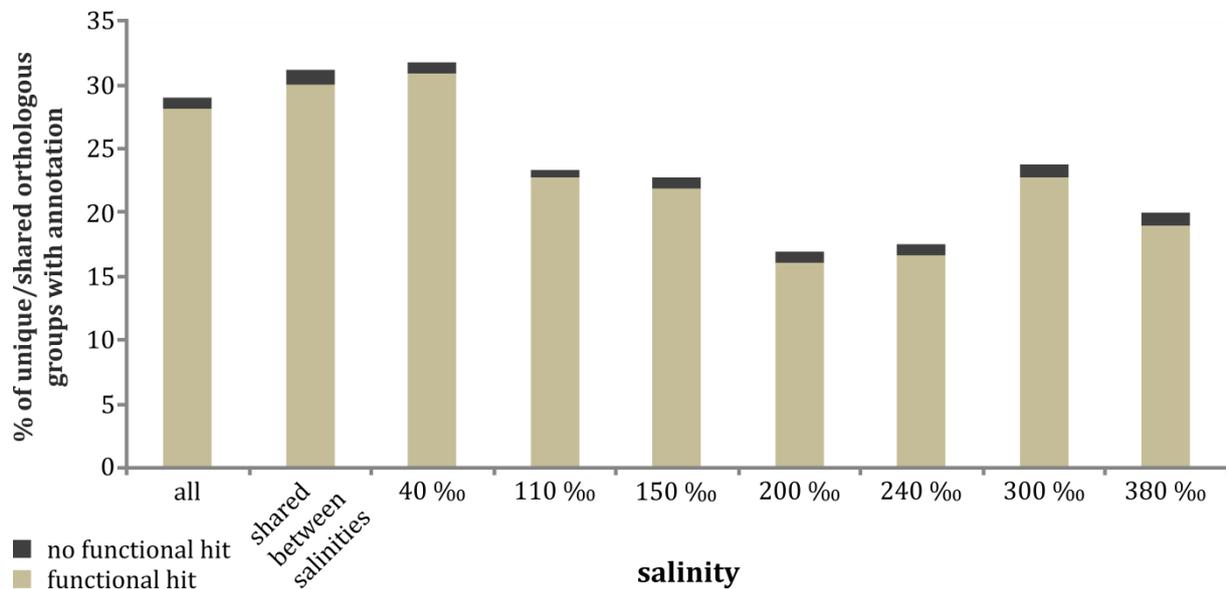
Figure 15: Proportion of annotated OGs according to salinity. On average, 28 % of the 1,165,559 OGs could be functionally annotated. The highest amount of functionally annotated OGs was found for shared OGs between salinities and for unique OGs at 40 ‰ salinity (30 % and 31 %). The proportion of annotated OGs decreased considering unique OGs of hypersaline samples. The group 'no functional hit' included annotations like 'unknown', 'predicted protein', 'hypothetical protein', 'unnamed protein' and 'uncharacterized protein'.

2.2 Taxonomic assignment of orthologous groups

The 1,165,559 OGs were assigned to taxonomic identities based on phylum level (see chapter '2.4.2 Annotation'). The relative proportion of each phylum per salinity was calculated using the transcripts per million (TPM) values resulting from the occurring OGs in the respective salinity. The largest amount of TPM was allocated to 'no hit' and 'multiple hit' in all salinities indicating that most OGs could not be reliably annotated. The proportion of unannotated OGs increased with increasing salinity (Fig. 16). Among the taxonomically annotated OGs appeared the dinophyceae, ciliophora and chlorophyta as the most dominant phyla throughout the different intvestigated samples. These dominant three phyla were complemented by haptophyceae in the 40 ‰ salinity sample and by rhizaria in the 110 and 150 ‰ salinity sample (Fig. 16). The dinophyceae constituted the most abundant phylum at a salinity of 40 and 110 ‰. A high proportion of ciliophora was observed in the samples of 110 and 150 ‰ salinity. At higher salinities, i. e. between 200 and 380 ‰ salinity, the microeukaryotic communities were dominated by chlorophyta (Fig. 16). The highest diversity in terms of phylum number was detected in the 40 ‰ salinity sample, whereas diversity showed a decreasing trend with increasing salinity (with exception of the 300 ‰ salinity sample; Fig. 16).
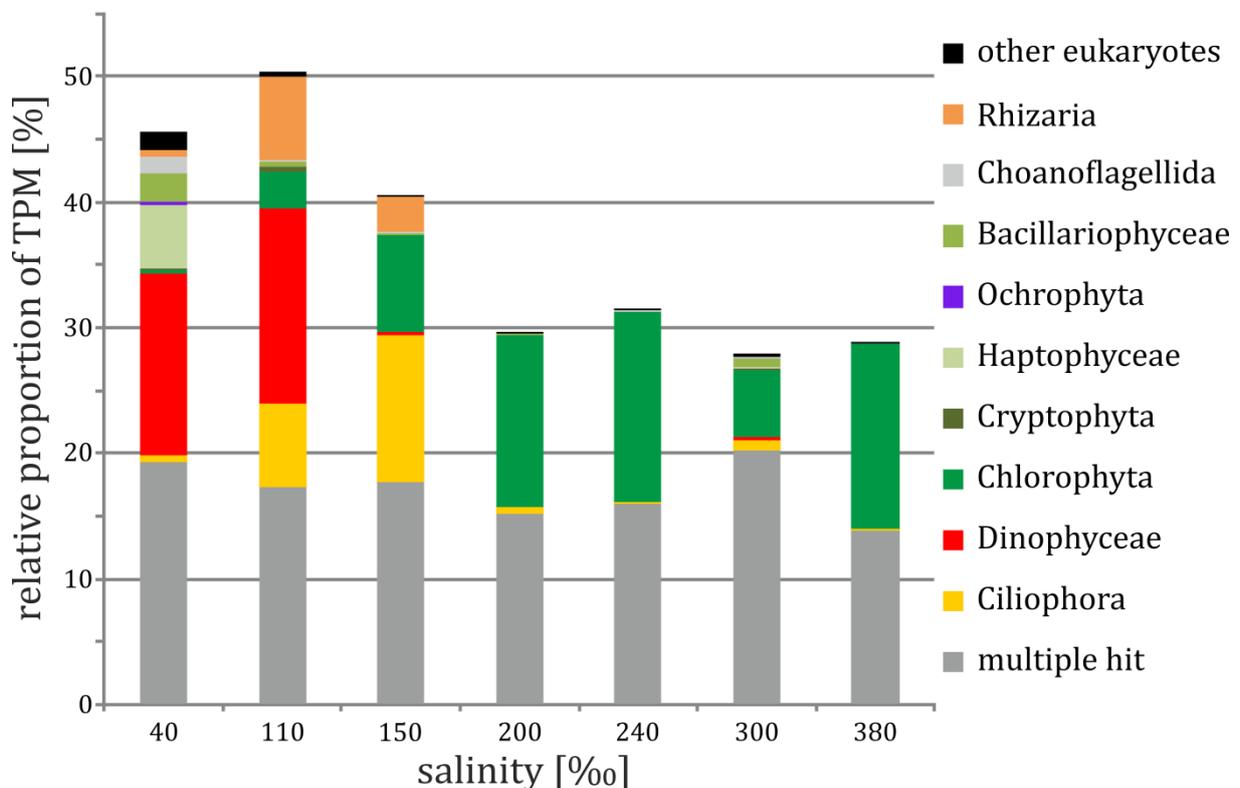
Figure 16: Transcript abundance-based phylum specific community composition on the basis of taxonomically assigned OGs with an e-value cut-off of $1e^{-5}$ as relative proportion of TPM according to salinity. Clear shifts in the major taxonomic groups among the different salinity samples could be

observed. Dinophyceae appeared as the most dominant phylum at 40 and 110 ‰ salinity, followed by ciliophora and rhizaria in salinities of 110 and 150 ‰. A large amount of haptophyceae could only be detected at 40 ‰ salinity. In hypersaline samples of 200 to 380 ‰ salinity, chlorophyta was the most abundant phylum. The greatest quantity of different phyla was observed at 40 ‰ salinity and the total number of different phyla decreased with increasing salinity reaching a high proportion of unassigned transcripts (gap to 100 % and 'multiple hits' ≥85 %) in high salinities (200 – 380 ‰). 'Multiple hits' contained taxonomic assigned OGs showing no consensus in their phylum identity.

2.3   Clustering of functional diversity

To investigate functional (dis)similarities between the samples, an UPGMA analysis was done and the samples from different salinities were clustered on the basis of their occurring OGs and their correspondend OG-abundance (see chapter '2.4.4 Partitioning of functional diversity'). All nodes separating the tree branches were well supported by bootstrap values of 1, indicating a robust tree topology (Fig. 17). Three main clusters could be observed: First, the sample with a salinity of 40 ‰ clustered completely separated from all other samples forming the first cluster, called 'marine' (40 ‰); second, the 110 ‰ and 150 ‰ salinity samples were grouped together forming the second cluster (mid-hypersaline); third, the samples with salinities ranging from 200 ‰ to 380 ‰ were markedly separated from the second cluster forming an extreme-hypersaline cluster (Fig. 17).
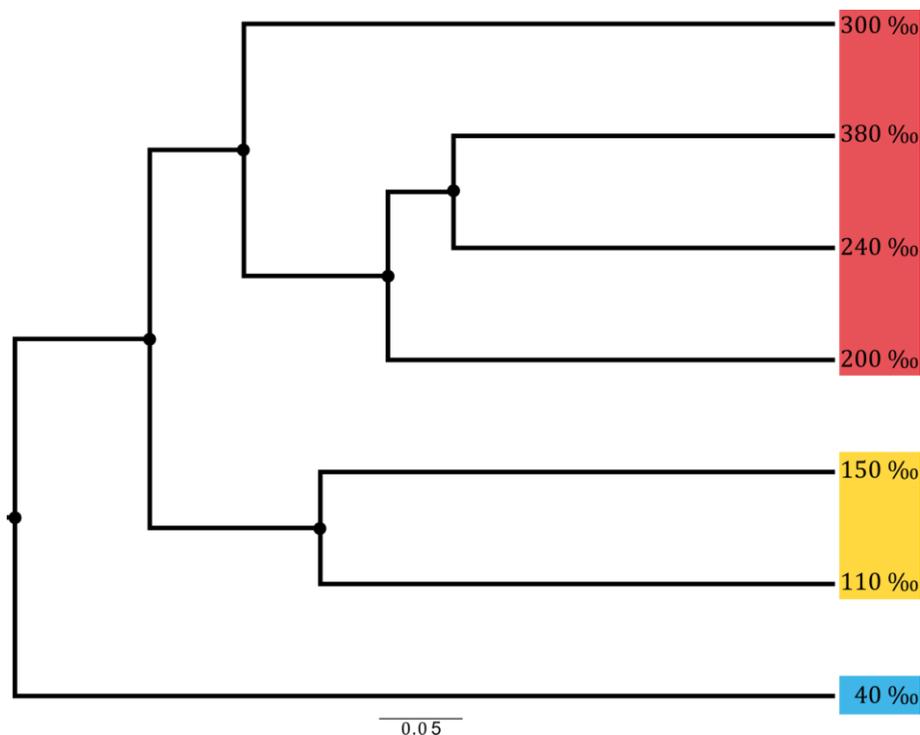
Figure 17: UPGMA clustering of Bray-Curtis based distance matrix, which is based on occurring OGs and their corresponding abundance in a sample. UPGMA clustering summarizes functional expression differences observed in the respective microeukaryotic communities of the different

salinities to show functional (dis)similarities. Three different clusters could be observed: a marine cluster (40 ‰; blue), a mid-hypersaline (110 – 150 ‰; yellow) and extreme-hypersaline (200 – 380 ‰; red) cluster. Bootstrap values indicate the strength of relationships in the dendrogram, with 1 (•) indicating a tree topology support of 100 %.

To identify relevant OGs or metabolic properties being responsible for the observed cluster pattern (Fig. 17), all OGs were mapped to the KEGG database (see chapter '2.4.2 Annotation'). OGs having a match in the KEGG database (n = 79,349) were sorted according to the KEGG classifications and module structure and together with their abundance represented in a heatmap (Fig. 18). The most abundant OGs throughout the different samples were classified in the 'nucleotide and amino acid metabolism', 'metabolism' and 'genetic information processing' module. The fewest OG abundances were found in the 'environmental information processing' module. A clear trend in expression value shifts among the different samples was generally not observed, instead, fluctuations occurred quite frequently. The OGs with the highest expression values were classified to the sub-module 'signal transduction' across all samples and to 'ribosome' especially at 40 ‰ salinity. The OGs with the lowest expression values were detected in the sub-module 'phosphotransferase system (PTS)' at 150 and 380 ‰ salinity (Fig. 18).

Since KEGG provides classifications on a rather low, i. e. general level, potential osmotically relevant strategies to tolerate different salinities were masked in the sub-modules 'two component regulatory system', 'signal transduction', 'carbohydrate metabolism', 'energy metabolism', 'serine and threonine metabolism', 'membrane transport' and 'other carbohydrate metabolism'. For these categories in specific no significant changes in expression connected to the changes in salinity could be observed (Fig. 18).
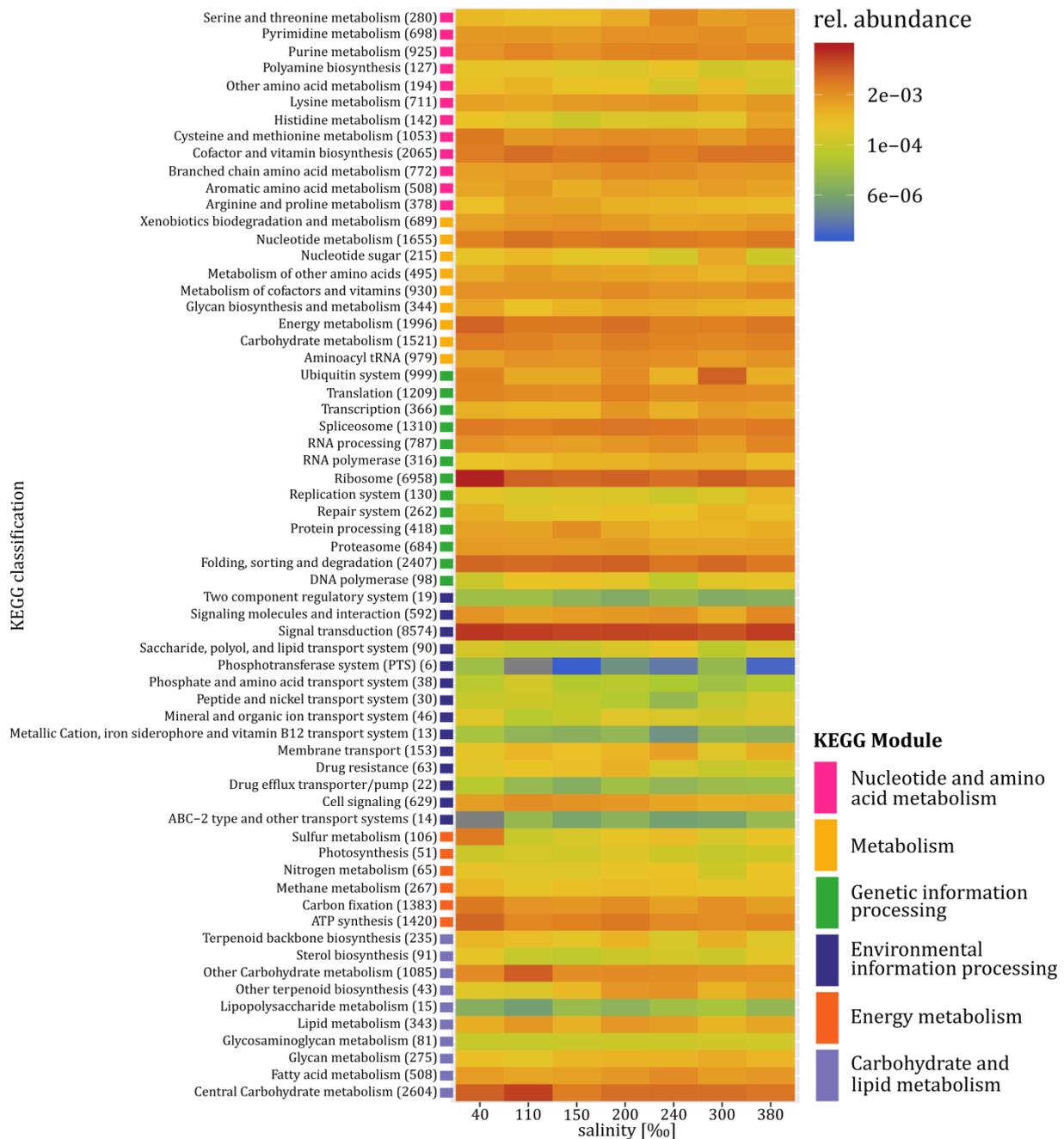
Figure 18: Heatmap of OG-based annotated functions based on the KEGG database. Only hits were considered with an e-value cut-off of 1e-3, which is automatically implemented in the GHOSTX algorithm (Suzuki et al., 2014). Observed functions could be assigned to the main KEGG modules 'nucleotide and amino acid metabolism', 'metabolism', 'genetic information processing', 'environmental information processing', 'energy metabolism' and 'carbohydrate and lipid metabolism'. The numbers in brackets following the functional annotation give the number of OGs assigned to a function. Relative abundances of annotated OGs appear to be relatively uniform in the different salinity samples, thus, not revealing significant patterns, which could explain the observed clustering pattern as depicted in Fig. 17.

Focusing only on osmotic relevant processes and intracellular reactions with a higher resolution, 44,045 differentially expressed OGs were extracted and manually grouped according to their annotated function into five defined categories: 'sensing',

'compatible solute transport', 'compatible solute metabolism', 'ion transport' or 'energy'. These datasets were then again subjected to UPGMA analyses and heatmap constructions in order to investigate whether these categories may explain the observed partitioning of functional diversity (Fig. 17).

2.3.1 Sensing

The UPGMA analysis based on all annotated functions belonging to the category 'sensing' revealed three distinct groups reflecting functional expression differences within community clusters according to salinity. The observed clustering pattern was congruent to the previously observed clustering pattern, presenting a marine cluster (40 ‰), mid-hypersaline cluster (110 – 150 ‰) and extreme-hypersaline cluster (200 – 380 ‰). All nodes of the dendrogram were supported by maximal bootstrap values of 1 (Fig. 19).
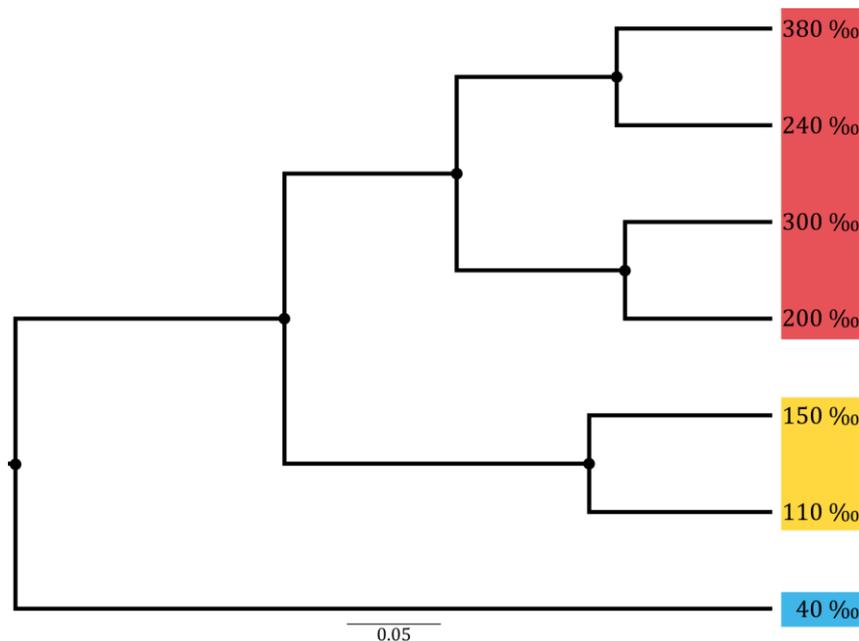
Figure 19: UPGMA clustering of Bray-Curtis index based distance matrix calculated using transcript abundances of functional annotations belonging to the category 'sensing'. The clustering revealed three distinct groups: a marine cluster (40 ‰; blue), a mid-hypersaline (110 – 150 ‰; yellow) and extreme-hypersaline (200 – 380 ‰; red) cluster. Bootstrap values indicate the strength of relationships in the dendrogram (• = 1). The different clusters are colour marked as it was shown before (Fig. 17).

In total, 969 OGs were manually assigned to the category 'sensing' representing 70 different functional groups. The 40 ‰ salinity sample differed clearly containing the

fewest portion of different functions. Most pathways were found at 110 and 150 ‰ and the amount of different functions decreased slightly for higher salinities again (Fig. 20). Exemplary, the following functions varied with their abundance referred to the three different groups and therefore, their capacity change between the three groups could be responsible for the cluster pattern of 'sensing'. First, the two component response showed the highest capacity for the mid-hypersaline cluster with decreasing capacity in higher salinities and missing completely for the marine cluster (40 ‰). Second, annotations of protein-coding gene regions belonging to stress response had the lowest capacity at 40 ‰, increasing constantly with increasing salinity until the highest capacity for protein-coding gene regions belonging to stress response was detected for communities in the extreme-hypersaline cluster (Fig. 20). Third, a similar pattern of their capacity was found for the thioredoxin-based redox signalling, starting low and increasing with increasing salinity (Fig. 20). Forth, annotated proteins of osmotic avoidance were only detected in the mid-hypersaline cluster just like cGMP dependent protein kinases (Fig. 20). Fifth, the only category of annotated proteins with a capacity distribution in the opposite direction was calmodulin having the highest capacity in the community of the marine cluster and decreases slightly with increasing salinity (Fig. 20). Finally, the highest number of different functional categories inside 'sensing' was found inside the mid-hypersaline cluster and the fewest were detected for the marine cluster. Additional examples are the different types of MAP-kinases and their related molecules and targets (Fig. 20).
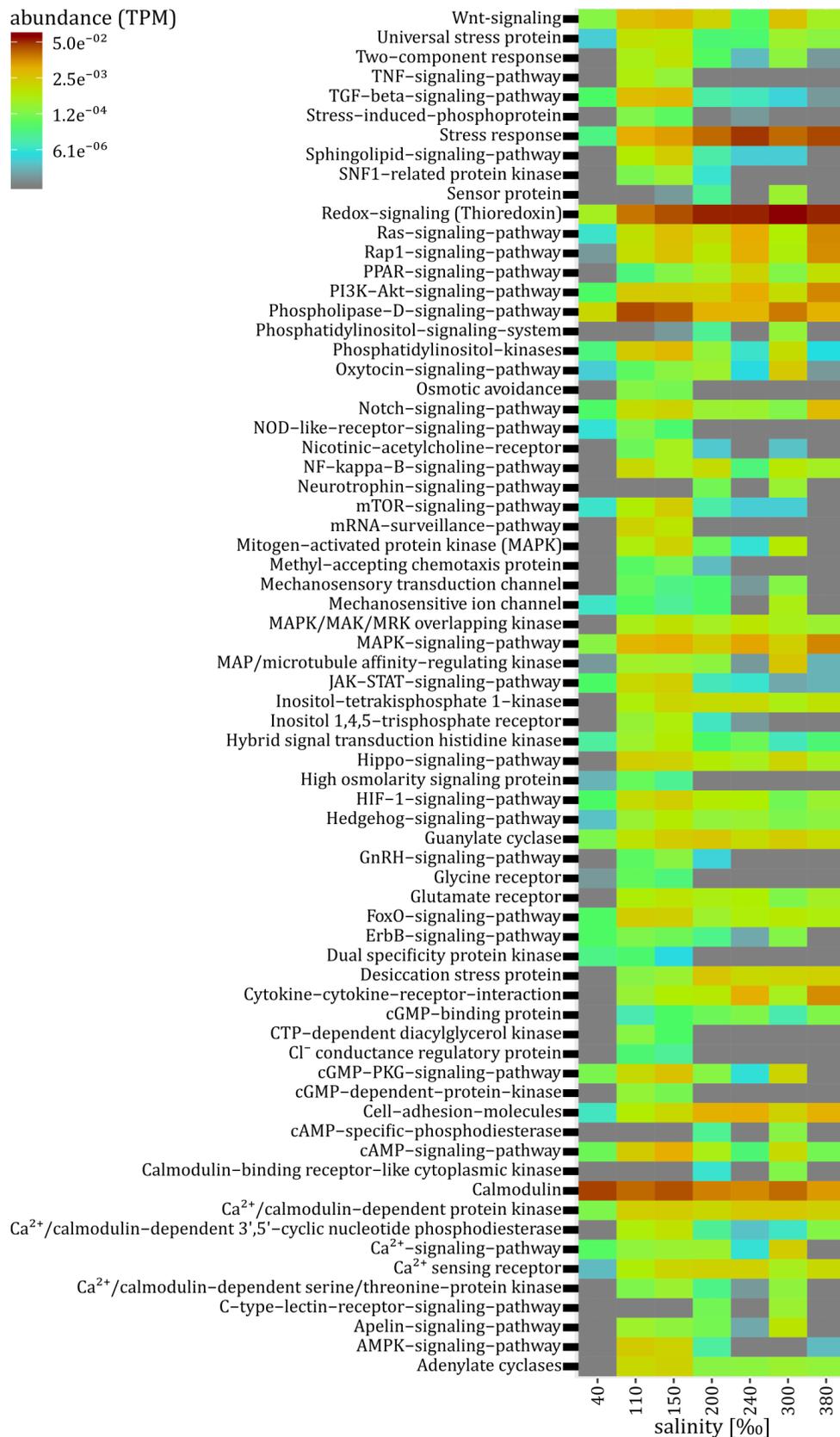
Figure 20: Heatmap of functional capacities within the microeukaryotic communities living at the different salinities being assigned to the category 'sensing'. Only OGs with differential expression between the certain salinity samples were considered for capacity calculation. Some exemplary functions, like two component response, stress response, redox signaling, phospholipase D signaling pathway, osmotic avoidance, mitogen activated protein kinases adenylate cyclases or calmodulin, reflected the observed cluster pattern of the UPGMA (Fig. 19).

2.3.2 Compatible solute transport

The annotated functions of the category 'compatible solute transport' provided with their UPGMA analysis a very similar cluster pattern as mentioned above (see Fig. 17) forming three distinct groups: marine, mid-hypersaline and extreme-hypersaline. Additionally, the separation of the marine cluster from the hypersaline samples was well supported by a maximal bootstrap value of 1. The partitioning of the mid-hypersaline cluster from the extreme-hypersaline cluster could be assumed, as 70 % of all calculated dendrograms supported this segregation (bootstrap value of 0.7; Fig. 21).
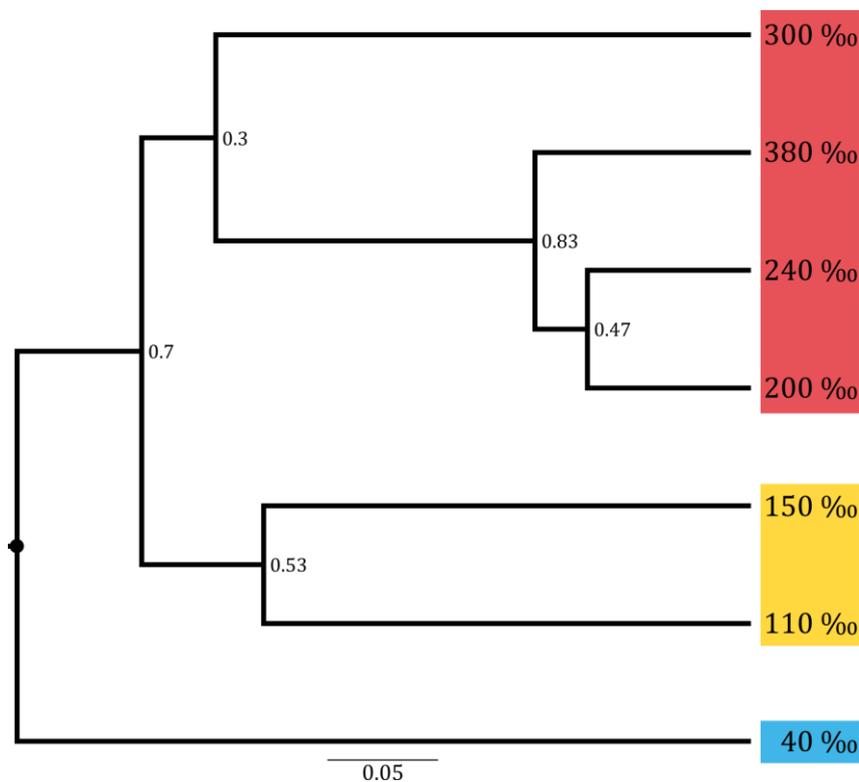
Figure 21: UPGMA clustering of Bray-Curtis index based distance matrix calculated using transcript abundances of functional annotations belonging to the category 'compatible solute transport'. Bootstrap values indicate the strength of relationships in the dendrogram (• = 1). Three different clusters were observed and colour marked as it was shown before (Fig. 17): a marine cluster (40 ‰; blue), a mid-hypersaline (110 – 150 ‰; yellow) and extreme-hypersaline (200 – 380 ‰; red) cluster.

After detailed view inside the annotated functions of the 113 OGs belonging to the category 'compatible solute transport', 26 different transport mechanisms for at least 21 potential compatible solutes were identified (Fig. 22). Most transporters were found within the 110 and 150 ‰ salinity samples and only few transport mechanisms with low abundances in the marine sample (40 ‰ salinity). For example, the capacity

for sugar transport increased from 40 ‰ to the mid-hypersaline cluster and reached its maximum in the extreme-hypersaline group (Fig. 22). Conversely, capacities for the transport of glycine betaine or ectoine were only detected in marine and mid-hypersaline samples but missing in the extreme-hypersaline cluster (Fig. 22). Choline as compatible solute and a potential precursor for the glycine betaine synthesis (Suppl Fig. 6) had a significantly increased transport capacity from marine to mid-hypersaline samples but a decreased choline transport capacity was found for the extreme-hypersaline cluster (Fig. 22). Similarly, the capacity for trehalose transport had its maximum in the mid-hypersaline cluster and only low capacities at 40 ‰ and salinities >200 ‰ (Fig. 22).
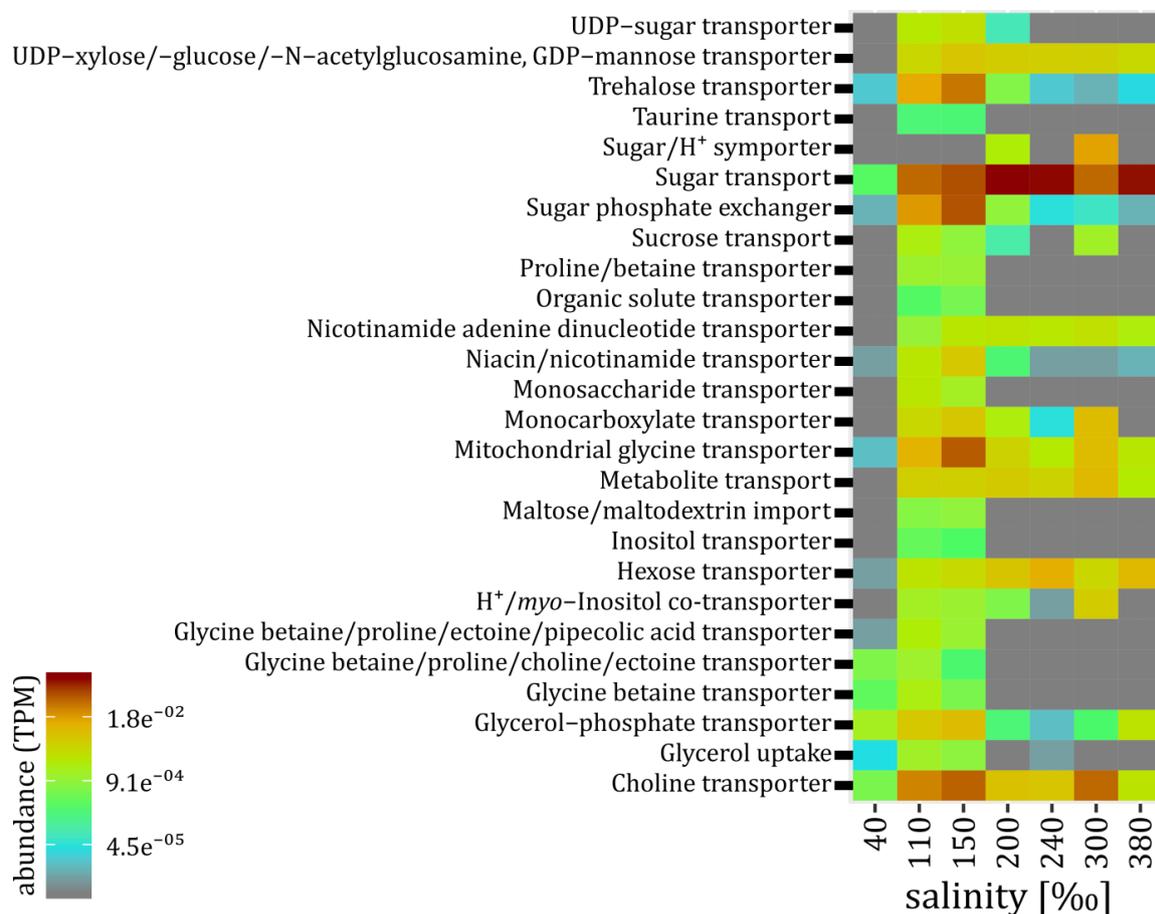
Figure 22: Heatmap of the different capacities for 'compatible solute transport' according to the different salinities. Only OGs with differential expression between the respective salinity samples were considered for capacity calculation. Some exemplary transporter of i. e. trehalose, taurine, sugar, glycine betaine, ectoine or choline reflected the observed cluster pattern of the UPGMA (Fig. 21).

2.3.3 Compatible solute metabolism

The annotated functions of the category 'compatible solute metabolism' originated from 373 OGs and contained enzymes for the *de novo* synthesis or the catabolism of 17 known compatible solutes, namely trehalose, taurine, sucrose, sarcosine, proline, *myo*-inositol, hydroxyectoine, hydroxybutyrate, glycine betaine, glycerol, glutamine amide, glutamate, ectoine, choline, aspartate, asparagines and alanine. Only by comparison of the synthesis and the catabolic capacity for a specific compatible solute, it is possible to assume whether it is more likely that a compatible solute was accumulated predominantly in this community by protists or mainly used as potential energy source (Fig. 24). The UPGMA cluster pattern of compatible solutes supported only the separation of marine communities from hypersaline ones (Fig. 23). All other nodes had very low bootstrap values and only the grouping of the 110 ‰ and 150 ‰ samples to form the mid-hypersaline cluster could still be assumed as the most likely tree is shown (Fig. 23).
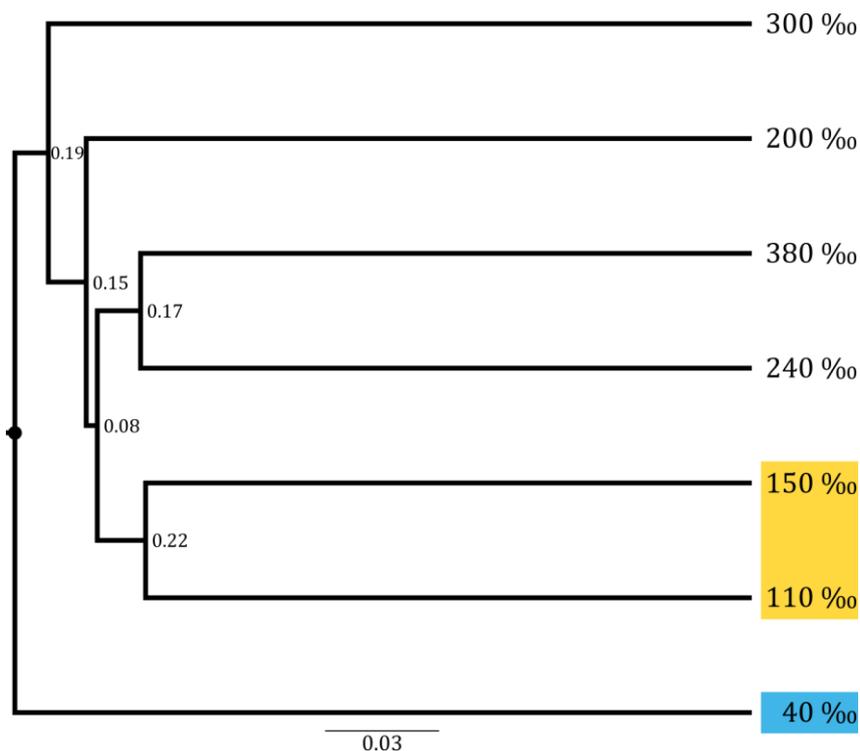
Figure 23: UPGMA clustering of Bray-Curtis index based distance matrix calculated using transcript abundances of functional annotations belonging to enzymes for the *de novo* synthesis or catabolism of compatible solutes. Bootstrap values indicate the strength of relationships in the dendrogram (• = 1) and supported only the separation of the marine cluster (40 ‰; blue) from hypersaline samples. The most likely tree is shown assuming a mid-hypersaline cluster is possible (110 – 150 ‰; yellow).

Comparing the capacities for the *de novo* synthesis and the catabolism of the 17 detected compatible solutes, the high capacity of both synthesis and catabolism of glutamate was significant in all hypersaline samples (Fig. 24). In contrast, the capacity for *myo*-inositol synthesis exceeded the capacity for the degradation in hypersaline communities (Fig. 24). Similar results were found for trehalose, sucrose and proline (Fig. 24). However, the catabolic capacity of glycerol and aspartate was significantly higher than for glycerol or aspartate synthesis at all salinities (Fig. 24). For some compatible solutes no degrading enzymes were detected. For example, an increased capacity for hydroxyectoine synthesis occurred only in the mid-hypersaline cluster while the capacity for glycine betaine synthesis increased with increasing salinity instead (Fig. 24). In case of asparagine, the received result was different: in the marine cluster, the capacity of enzymes for asparagine synthesis outperformed their catabolic counterparts, but in mid-hypersaline communities the capacity for asparagine synthesis and catabolism was balanced while the synthesis capacity was overbalanced in extreme-hypersaline samples again (Fig. 24). In contrast, the capacity patterns for choline synthesis or catabolism as well as for ectoine synthesis and catabolism revealed no clear pattern according to different salinities indicating no distinct relationship to the observed cluster pattern (see Fig. 23).
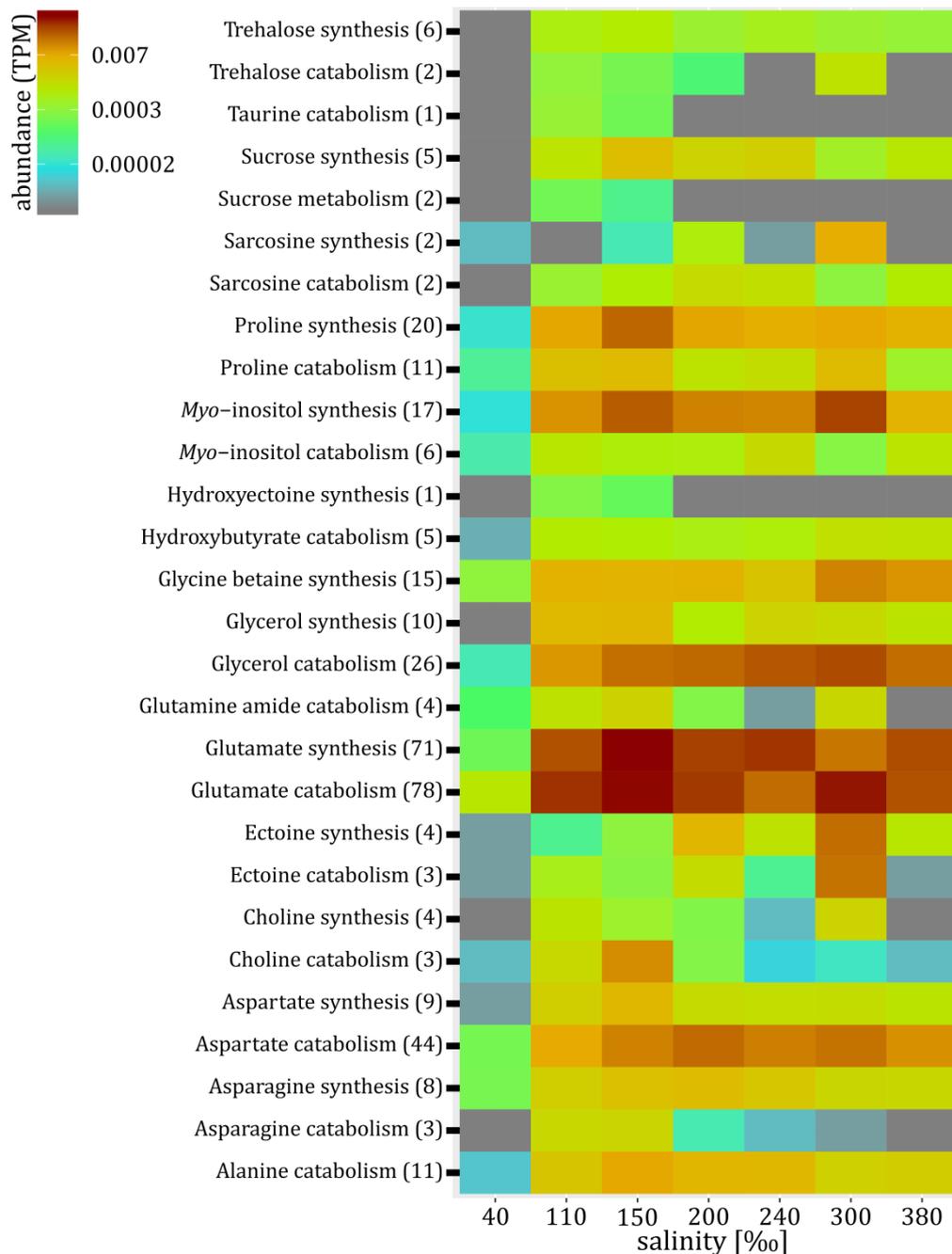
Figure 24: Heatmap of the different capacities for the synthesis and catabolism of known compatible solutes according to different salinities. Only OGs with differential expression between the respective salinity samples were considered for capacity calculation and their numbers are shown in brackets. For some compatible solutes at certain salinities, the synthesis capacity was greater than the catabolic counterpart. For others at different salinities, the catabolic capacity exceeded the synthesis capacity or both were balanced.

2.3.4 Ion transport

The annotated functions regarding mechanisms for ion transport demonstrated with their cluster pattern of the UPGMA analysis a distinct separation of the 40 ‰ from all other hypersaline samples (Fig. 25). The relatively low bootstrap values between

hypersaline samples indicated an unreliable positioning of the 300 ‰ sample. Only 51 % of all calculated trees placed this sample (300 ‰) next to the mid-hypersaline cluster reasoning that other groupings may be possible (Fig. 25). The high bootstrap value of 0.96 for the grouping of 110 ‰ and 150 ‰ indicates the existence of a mid-hypersaline cluster. An extreme-hypersaline cluster could not be identified because of the precarious positioning of 300 ‰ (Fig. 25).
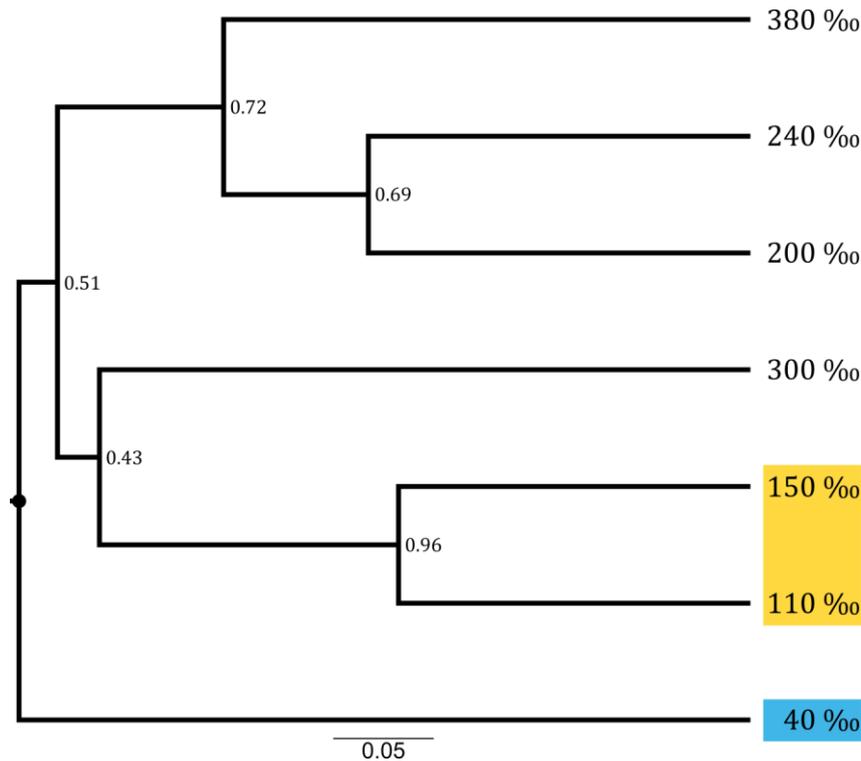
Figure 25: UPGMA clustering of Bray-Curtis index based distance matrix calculated using transcript abundances of functional annotations belonging to mechanisms for ion transport. Bootstrap values indicate the strength of relationships in the dendrogram (• = 1) and supported only the separation of the marine cluster (40 ‰; blue) from hypersaline samples. The most likely tree let appear a mid-hypersaline cluster very assumable (110 – 150 ‰; yellow) but the positioning of 300 ‰ is unsecure avoiding the observation of a extreme-hypersaline cluster.

The 336 manually grouped OGs into the category 'ion transport' represented 46 different ion transport mechanisms according to their annotated functions. Only 15 different ion transport mechanisms were found within the 40 ‰ salinity sample showing the lowest abundances. In contrast, the detected mechanisms for ion transport provided the greatest transport capacities for inorganic ions in samples of the mid-hypersaline cluster (110 – 150 ‰ salinity). The amount of different ion transport mechanisms decreased slightly from the mid-hypersaline cluster to the extreme-hypersaline cluster (> 200 ‰) but outperformed the marine cluster significantly (Fig.

26). For example, the capacity of voltage-gated K^+ channels is relatively low at 40 ‰ and higher in hypersaline samples with its maximum at 110 ‰ and 150 ‰ (Fig. 26). Similar observations could be made for the capacity of Na^+/K^+-transporting ATPases and Na^+/H^+ antiporter as well as for cation-transporting ATPases, Ca^{2+}-transporing ATPases and Ca^{2+}-activated K^+ channels (Fig. 26). Different capacity patterns according to the salinity of the certain samples demonstrated the Mg^{2+} transport. The transport capacity for Mg^{2+} had no significant difference investigating the marine group (40 ‰ salinity; grey colour; Fig. 26) but the capacity increased with increasing salinity in hypersaline samples reaching their highest amount within the extreme-hypersaline group (200 – 380 ‰; Fig. 26). Also, untypical capacity pattern was observed for annotations of general ion transporter. Their capacity increased with increasing salinity until attaining its maximum at 240 ‰, showing the highest capacity of all functional annotated mechanisms for ion transport (Fig. 26). A constantly relatively high capacity, instead, was visible for anion transporter independent of the salinity of the sample (Fig. 26). Furthermore, a lot of other transport mechanisms showed only an increased transport capacity with significant difference compared to other salinities for the mid-hypersaline cluster by grey coloured fields for 40 ‰ and extreme-hypersaline samples in contrast to varicoloured fields of 110 ‰ and 150 ‰ (Fig. 26). Examples could be found for voltage-gated Na^+ channels, voltage-dependent Ca^{2+} channels, vacuolar cation-transporting ATPases, Na^+/pyruvate co-transporter, Na^+/pantothenate symporter, Na^+/bile acid co-transporter, H^+/Cl^- exchanger, cys-loop ligand-gated cation channels, cGMP-gated cation channels, cation/Cl^- co-transporters, cation efflux and Ca^{2+} permeable stress-gated cation channels (Fig. 26).
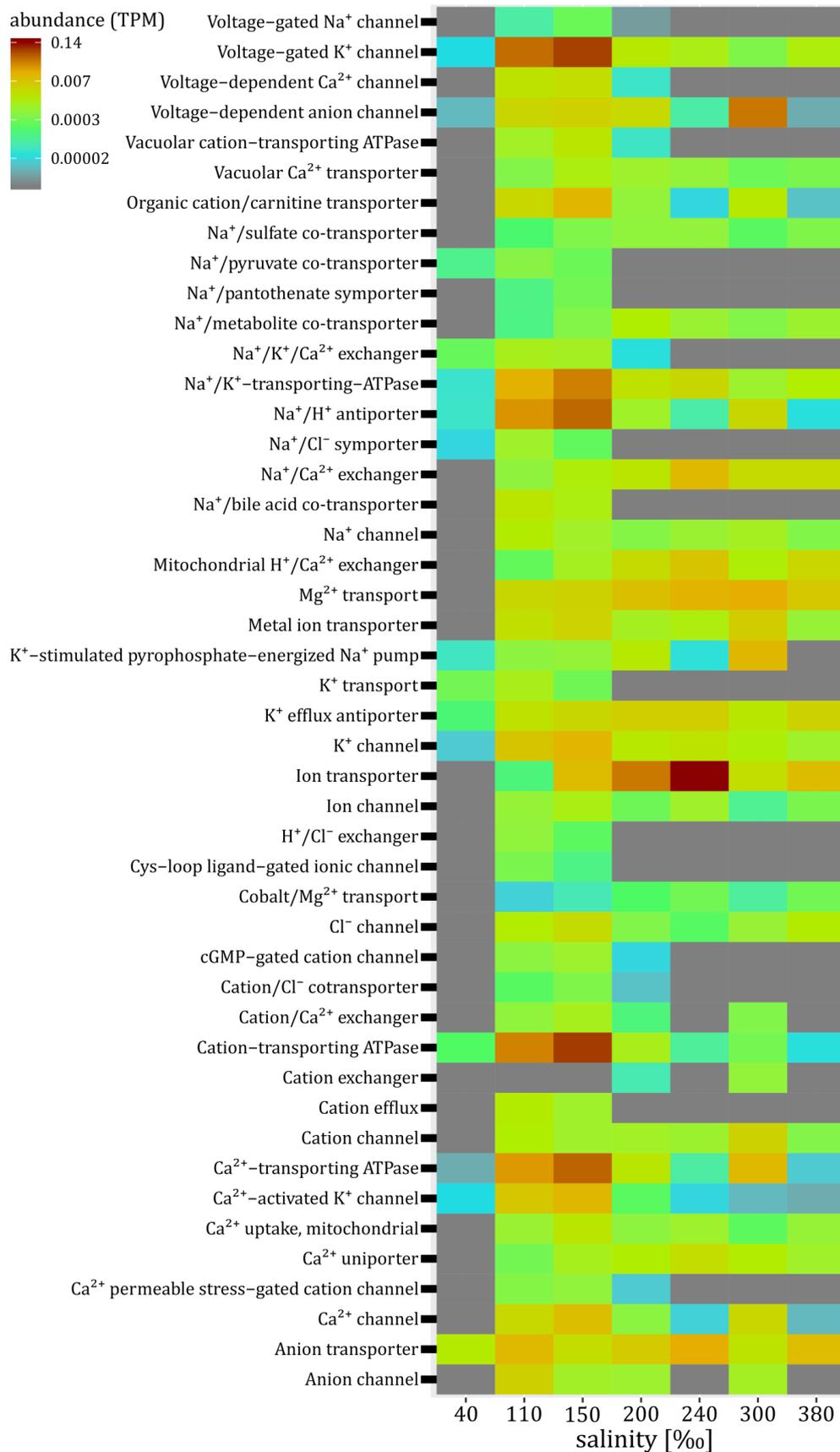
Figure 26: Heatmap of the different capacities for ion transport mechanisms according to the different salinities of the samples. Only OGs with differential expression between the respective salinity samples were considered for capacity calculation. For most transport mechanisms, the highest capacity was found within the mid-hypersaline cluster (110 – 150 ‰).

2.3.5 Energy

The annotated functions of the category 'energy' provided with their UPGMA analysis a completely different result as the dendrograms shown above. Strong bootstrap values supported a clear separation of 40 ‰ salinity sample from all other samples forming the marine cluster (Fig. 27). All other hypersaline samples could be considered together to form one big hypersaline cluster. A detailed separation within this hypersaline group was not observed and not supported by bootstrap values (Fig. 27).
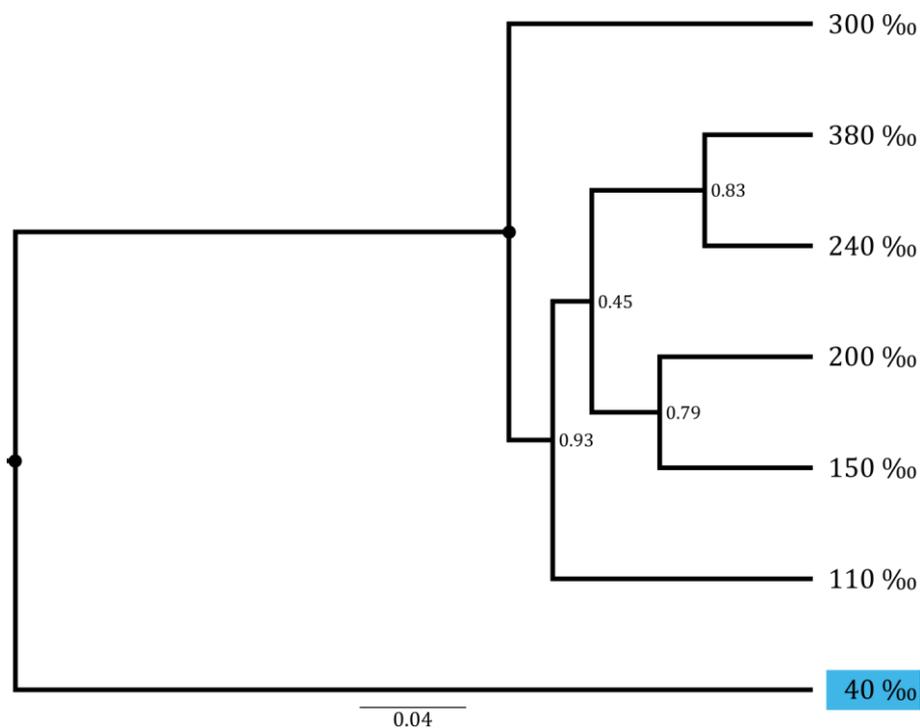
Figure 27: UPGMA clustering of Bray-Curtis index based distance matrix calculated using transcript abundances of annotated functions of the category 'energy'. Bootstrap values indicate the strength of relationships in the dendrogram (• = 1) and supported only the separation of the marine cluster (40 ‰; blue) from hypersaline samples.

With a detailed view inside the annotated functions of 3691 OGs belonging to the category 'energy', 41 different mechanisms or energy-related processes were detected. Most of them were found within the hypersaline samples (110 – 380 ‰) and only 31 within the marine cluster (40 ‰ salinity). Especially, the capacities of the energy-related processes differed significantly between hypersaline samples and 40 ‰ (Fig. 28). With the exception of sulfur metabolism, all other investigated processes had an increased capacity in hypersaline samples and a lower capacity at 40 ‰ salinity, for example 'V-type H$^+$ ATPase', 'steroid biosynthesis', 'starch biosynthesis',

'photorespiration', 'nitrogen metabolism', 'lipid biosynthesis', 'GMP synthase', 'glycogen catabolism', 'fatty acid metabolism', 'fatty acid biosynthesis', 'coenzyme A biosynthesis', 'citric acid cycle (TCA cycle)', 'chlorophyll biosynthesis', 'carotenoid biosynthesis', 'carbon fixation', 'calvin cycle', 'ATPase' or 'alcohol metabolism' (Fig. 28). Particularly, the capacity of lipid metabolism increased constantly with increasing salinity reaching its maximum at 380 ‰ salinity (Fig. 28). According to their constantly high capacities, the most dominant processes in hypersaline samples were 'respiration', 'photosynthesis', 'light harvesting', 'glycolysis' and with a slightly decreased capacity 'carbohydrate metabolism' as well 'ATP synthase' (Fig. 28). Variations within the hypersaline samples to separate the mid-hypersaline cluster (110 – 150 ‰) from the extreme-hypersaline cluster (>200 ‰) were really scarce and only found for 'xanthorhodopsin', 'glycogen synthesis regulation', 'glycogen biosynthesis', 'creatine metabolism' and 'arachidonic acid metabolism' (Fig. 28). A more detailed resolution of energy-related processes to find more significant variations within the hypersaline group was not possible. Hence, only higher capacities were clearly found for energy-related processes and energy production in hypersaline samples compared to the marine group (Fig. 28).
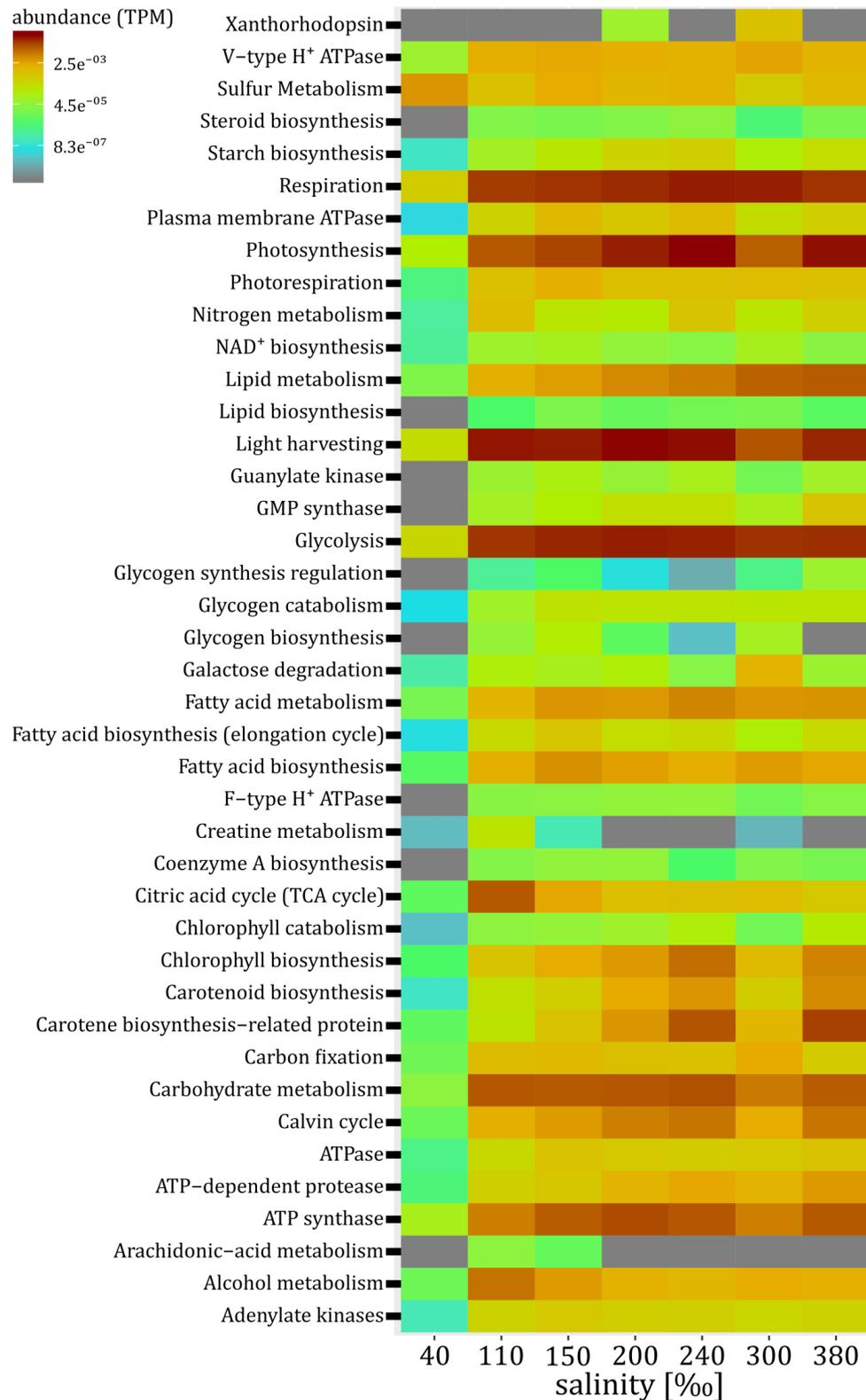
Figure 28: Heatmap of the different capacities of energy-related processes according to different salinities. Only OGs with differential expression between the respective salinity samples were considered for capacity calculation. In general, a higher capacity was found for hypersaline samples compared to the marine cluster (40 ‰). Further variations within the hypersaline group (110 – 380 ‰) were not possible to detect because of the basic necessity of energy production for living organisms and the ubiquitous distribution of their related processes. This prevented a more specific resolution of energy-related cellular processes showing broader categories like 'respiration', 'photosynthesis', 'lipid metabolism', 'light harvesting', 'glycolysis' or 'carbohydrate metabolism'.

2.4 Isoelectric points of cytoplasmic proteins

To detect a potential change of the used haloadaptation strategy within a protistan community besides measuring and counting compatible solutes or transport capacities, the isoelectric points of cytoplasmic proteins were compared between the different salinities. If the majority of proteins represented more acidic values for their isoelectric point, these proteins were less sensitive to high inorganic ion concentrations. Hence, assuming that the proteins of a certain salinity reflected with their isoelectric points the average acidity of the cytoplasma of the whole community living at this salinity, it provided an additionally hint whether the majority of the community has a well-adapted cytoplasm for tolerating higher intracellular ion concentrations or not. Comparing the calculated average acidity for all proteins between the seven salinity samples, no distinctive shifts in averaged acidity of the protistan communities could be observed (Fig. 29). Eventually, all isoelectric points of the different proteins were placed within a 95 %-confidence interval and these 95 %-confidence intervals were overlapping by comparing the different salinities with each other (borders of the boxplots in Fig. 29). Nevertheless, statistical differences were calculated on the basis of the mean of the isoelectric point of all occurring proteins (letters in Fig. 29). An absent shift in averaged acidity was also found when proteins of the different communities were separated according to their taxonomic affiliation or the three salinity cluster (marine, mid-hypersaline and extreme-hypersaline) within in the same phylum (Suppl. Fig. 7).
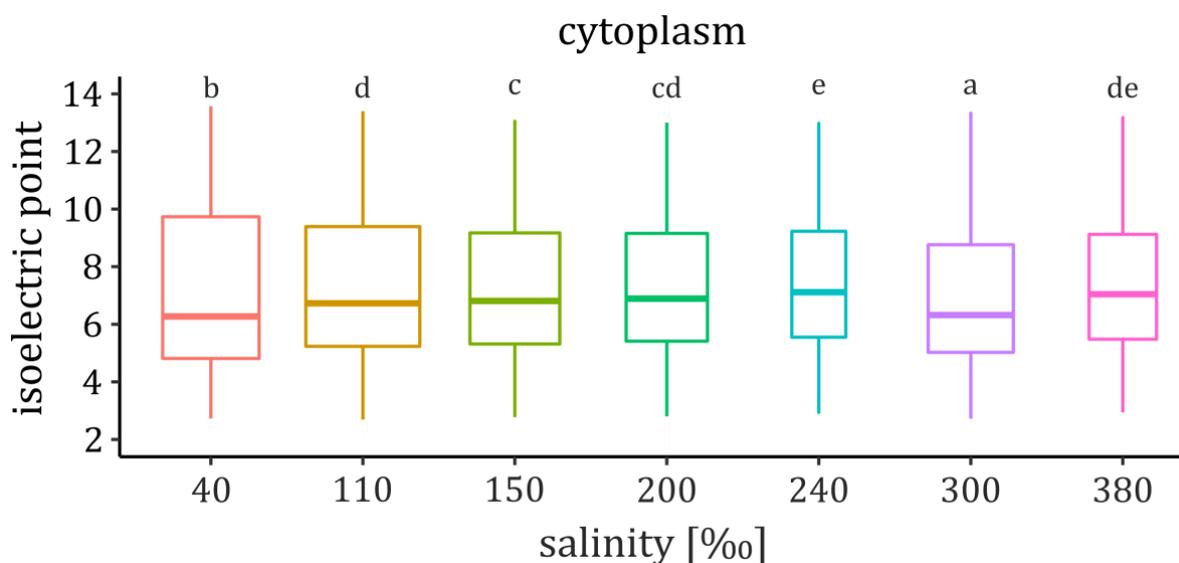
Figure 29: Comparing illustration of the IPC results according to different salinities. All boxplots are overlapping each other and the width was related to the amount of proteins used for the calculation. Statistical differences of the mean are symbolised by the usage of different letters (p < 0.05)

3. Time-resolved transcriptome analysis of *Schmidingerothrix salinarum*

As revealed by growth experiments, *Schmidingerothrix salinarum* is able to tolerate salinities between 38 and 220 ‰ and has a growth optimum at around 90 ‰ under lab conditions (Weinisch et al., 2018b). Additionally, using ^{1}H-NMR spectroscopy, it was shown that the ciliate is able to accumulate the compatible solutes glycine betaine and ectoine to counterbalance osmotic stress (Weinisch et al., 2018b), thus, being a 'low-salt – organic-solutes-in' strategist. The applied ^{1}H-NMR spectroscopy approach, however, did not provide information about involved proteins, enzymes, specific pathways and cellular processes associated with the 'low-salt – organic-solutes-in' strategy. In general, knowledge on respective pathways for the synthesis of glycine betaine in ciliates and ectoine in microeukaryotes was missing. Due to the existing knowledge on the ecophysiology of *S. salinarum*, the species was used as model organism to investigate the 'low-salt – organic-solutes-in' strategy in more detail, which was hypothesized to be the preferred osmoadaptation strategy used by protists thriving in low- to mid-hypersaline conditions (cf. subject 'II'). For this reason, *S. salinarum* was cultivated at 38 ‰ salinity and then subjected to a salt-up shock to 90 ‰ salinity. After specific timepoints, subsamples of the culture were taken to monitor the cellular expression dynamics after osmotic up-shock and during salt acclimation applying a time-resolved transcriptome analysis.

3.1 Sequence data and transcriptome overview

After sequencing of the 21 samples (seven timepoints, three replicates each), the obtained 28,158,492 to 69,817,402 reads per replicate and time point were reduced by means of adapter trimming, quality control and removal of ribosomal transcripts to 16,758,445 – 34,975,358 trimmed and sorted reads per replicate and time point (Suppl. Tab. 5). In total 498,216,987 reads were used as input for *de novo* transcriptome assemblies. Since the global transcriptome assembly comprising read data of all samples returned better assembly statistics (Tab. 7), the global assembly was used for ORF prediction and differential expression analyses.

Table 7: Contig metrics and assembly statistics sorted by the different *de novo* assemblies created with the Trinity pipeline (see chapter '3.4.1 Time-dependent transcriptome processing'). The number of raw reads was received after sequencing and the number of usable reads was used as input for assembly preparation. Contigs were characterized by the average read length and total number of bases. The assembly statistics provide detailed information about the quality of an assembly. The assembly score is calculated as the geometric mean of all contig scores multiplied by the proportion of input reads that provide positive support for the assembly (Smith-Unna et al., 2016). The optimal score resulted from the former assembly score by filtering out contigs with bad contigs scores (Smith-Unna et al., 2016). A higher assembly score defines a better assembly because an increased score is very likely to correspond to an assembly that is biologically more accurate (cf. Smith-Unna et al., 2016).

Contig metrics	global	0 min	2 min	10 min	30 min	60 min	120 min	720 min
no. of raw reads [million]	821.1	58.7-69.8	28.2-31.7	33.7-36.7	31.9-38.9	32.1-35.4	30.2-38.9	42.3-43.3
no. of usable reads [million]	498.2	31.8-34.9	16.8-19.8	22.0-23.0	21.2-25.8	20.5-22.7	17.8-23.7	24.5-28.2
average read length [bp]	749	639	557	632	854	678	546	753
total no. of bases [million]	232.2	112.6	64.9	80.8	70.7	69.6	73.1	80.1

Assembly statistics

no. of contigs	310,249	176,330	116,374	127,991	82,803	102,630	134,010	106,407
N50	1,280	949	722	895	1,348	993	701	1,199
GC content [%]	43.28	45.74	40.01	41.28	36.63	38.00	41.66	39.68
assembly score [0-1]	0.48	0.19	0.23	0.26	0.22	0.22	0.21	0.20
optimal score [0-1]	0.52	0.23	0.29	0.32	0.31	0.30	0.25	0.27

After removal of potential prokaryotic contaminations, the global dataset contained 168,454 ORFs originating probably from *S. salinarum*, of which 15,093 were differentially expressed with a FDR < 0.001 and at least 4-fold-change ($\log_2$FC $\geq$ 2) compared to the control t0 (Tab. 8). After mapping the ORF sequences against different available databases, 49.6 % of all ORFs and 77.0 % of all differentially expressed ORFs could be annotated (Tab. 8). To investigate how many different genes are present in the genome of *S. salinarum* and to measure how many different genes were differentially expressed over all six time points, the total number of ORFs was not suitable due to the well documented number of gene isoforms and varying copy numbers in the phylum ciliophora justified by their nuclear dualism (Herrick, 1994; Prescott, 1994). Therefore, the amount of different annotated functions from all ORFs was counted yielding 48,022 uniquely annotated protein functions for the global dataset and 9,239 uniquely annotated protein functions for ORFs with an at least 4-fold-change ($\log_2$FC $\geq$ 2) expression (Tab. 8).

Table 8: Numbers of annotated ORFs according to the used reference database, separated according to the global dataset, and to the subset of ORFs with $\geq$ 4-fold-change.

database	global dataset (168,454 ORFs)		ORFs with $\geq$ 4-fold-change (15,093 ORFs)	
KEGG	21,861	(13.0 %)	3,202	(21.2 %)
Genbank	4,419	(2.6 %)	1,075	(7.1 %)
NCBI´s non-redundant nucleotide (nt) reference database	6,109	(3.6 %)	1,423	(9.4 %)
NCBI´s non-redundant protein (nr) reference database	48,966	(29.1 %)	11,479	(76.1 %)
Swissprot	63,663	(37.8 %)	4,630	(30.7 %)
UniProt	63,509	(37.7 %)	4,627	(30.7 %)
TCDB	19,845	(11.8 %)	2,640	(17.5 %)
TransportDB2	7,301	(4.3 %)	969	(6.4 %)
annotated ORFs	83,633	(49.6 %)	11,618	(77.0 %)
ORFs without annotation	84,821	(50.4 %)	3,475	(23.0 %)
unique annotated functions	48,022		9,239	

Since no completely sequenced reference genome of *S. salinarum* was available, the exact genome size and the number of genes were unknown. The completeness of the individual and global transcriptomes was therefore assessed by searching the KEGG

annotations for essential metabolic pathways and structural complexes (Beisser et al., 2017). Respective modules are a collection of defined functional units, which require that all enzymes necessary for the reaction steps or proteins constituting a complex are present. These pathways and structural complexes covered modules, which are relevant for the functioning of the cell, such as central carbohydrate metabolism, fatty acid metabolism, nucleotide metabolism, ATP synthesis, DNA/RNA polymerase, replication system, repair system, spliceosome, RNA processing, ribosome, proteasome, ubiquitin system and protein processing. Nineteen of the 21 pathways/complexes were found to be complete for the individual and the global transcriptomes (Fig. 30). Although two enzymes were missing for the pentose phosphate pathway and the fatty acid biosynthesis, respectively, the central modules appeared to be complete.
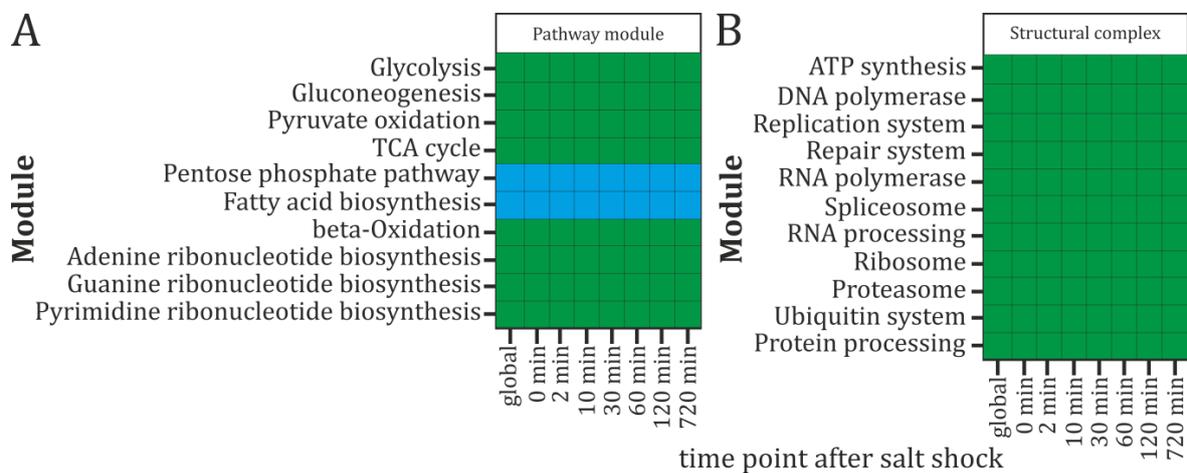
Figure 30: Completeness of KEGG essential modules following the methodology of Beisser et al. (2017). Modules were considered operational if all enzymes necessary for the reaction steps or proteins constituting a complex were present. Pathway modules (**A**) were coloured in green if at most one enzyme was missing, in blue if at most two enzymes were missing, but the central module was complete (e. g., complete module: M00001 Glycolysis (Embden-Meyerhof pathway); central module: M00002 Glycolysis, core module involving three-carbon compounds; cf. Beisser et al., 2017). Structural complexes (**B**) consisted of several modules and were coloured in green if all associated modules were present and at most one enzyme was missing per module but the module was still functional and all proteins constituting a complex were present. The modules and complexes were ordered according to the different assemblies and time points after salt shock shown on the x-axis. Supplemental data are given in the digital appendix in file 'Input_data_Fig.30_completeness_KEGG'.

To detect any differences between the samples and to confirm the congruence of all three replicates of a time point, the 21 samples were clustered according gene expression. All replicates of one time point were more similar to each other than to any other replicate from a different time point (Fig. 31). Furthermore, a cluster pattern for the different time points could be observed. First, the control group (t0) possessed the

greatest discrepancy from all other time points (Fig. 31). Second, the similarity of the six other time points formed two different groups. The first group consisted of the samples 2 min, 10 min, 60 min and 120 min after salt up-shock (Fig. 31). The other group contained the time points 30 min and 720 min after salt up-shock (Fig. 31).
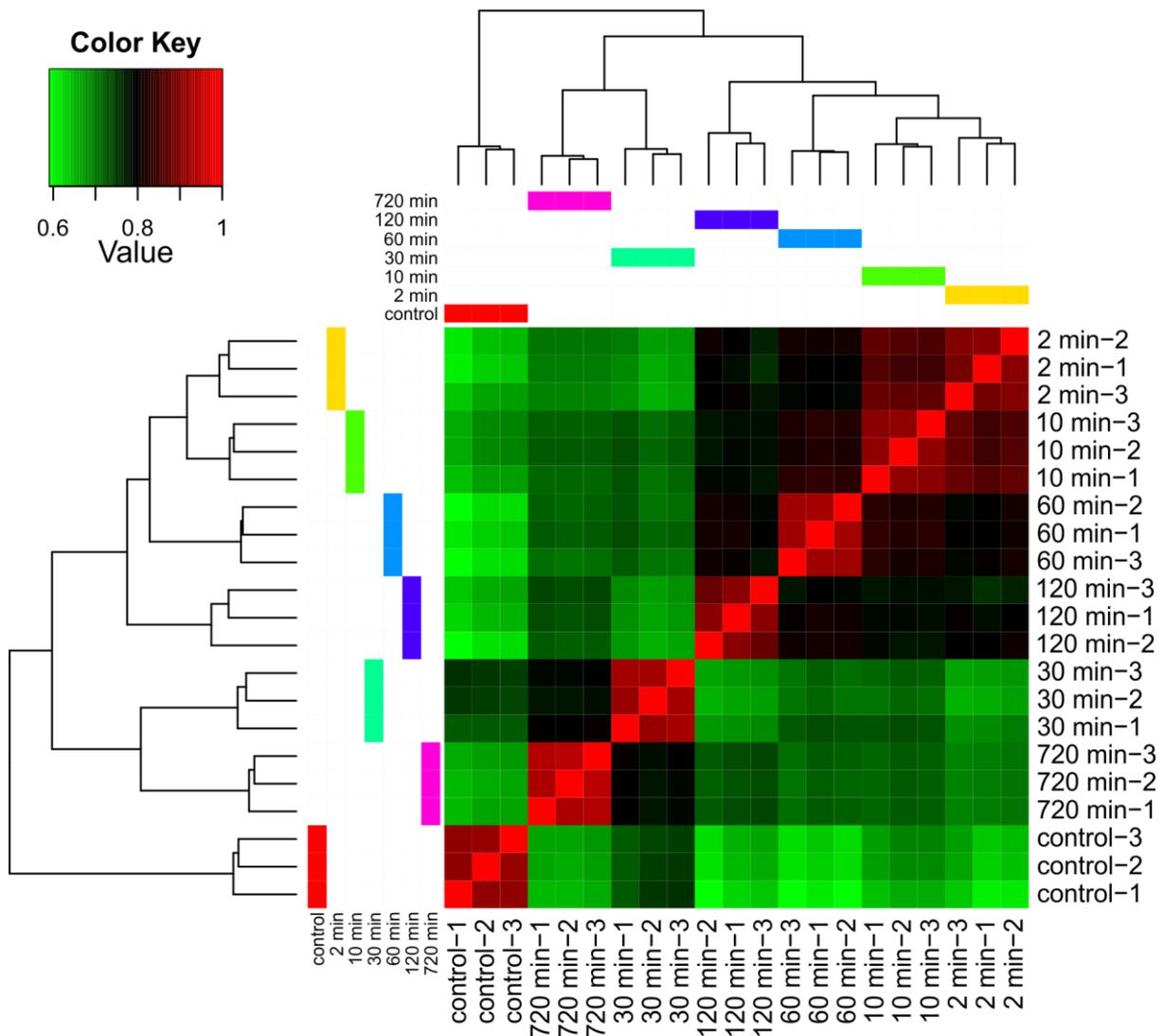
Figure 31: Sample correlation heatmap of replicates and time points according to their TMM-normalized expression counts per ORF. Only ORFs with a minimum expression of 0.1 in at least two samples were counted. A similarity value of 1 represented total equality of the two compared samples. With a value of 0, two compared samples were completely different. Within one time point, all replicates were more similar to each other than to any other replicate of a different time point (coloured bars). The sample clustering was based on Pearson correlation using default settings. Additionally, the separation of the control samples as well as a clustering of two groups could be observed: The first group contained all samples from 2 min, 10 min, 60 min and 120 min. The second group consisted of the replicates after 30 min and 720 min.

3.2   Time-dependent pattern of metabolic functions

Applying of a time-resolved transcriptome analysis allowed the investigation of gene expression pattern in *S. salinarum* cells at selected time points. Identified ORFs were considered differentially expressed, if the expression value of an ORF at a specific timepoint compared to timepoint zero (t0, right before salt-up shock) was $log_2FC \geq 2$ (up-regulated) or $log_2FC \leq -2$ (down-regulated). In total, 15,093 ORFs appeared to be differentially expressed and the number of differentially expressed ORFs varied over time (Tab. 9). After 2 min, 945 ORFs were up-regulated with at least a 4-fold-change ($log_2FC \geq 2$) and 540 ORFs were simultaneously down-regulated ($log_2FC \leq -2$). After 10 min, the number of down-regulated ORFs increased slightly to 579 and of up-regulated ORFs markedly to 2038 (Tab. 9). The highest quantity of up-regulated ORFs (5152) was found at 30 min after salt up-shock and simultaneously 1508 ORFs were down-regulated (Tab. 9). After 60 min, 2129 ORFs were up- and 622 ORFs were down-regulated (Tab. 9). The lowest amount of down-regulated ORFs (172) was observed at 120 min and at the same time 3425 ORFs were up-regulated (Tab. 9). After 720 min, 3190 ORFs were up- and 2108 ORFs were significantly down-regulated (Tab. 9).

Table 9: Number of up- and down-regulated differentially expressed ORFs (n = 15,093) for each time point after salt up-shock. Expression values of the individual time point were compared with the values of the control sample (t0).

time point after salt up-shock	ORFs with $log_2FC \geq 2$ (up-regulated)	ORFs with $log_2FC \leq -2$ (down-regulated)
2 min	945	540
10 min	2038	579
30 min	5152	1508
60 min	2129	622
120 min	3425	172
720 min	3190	2108

The differentially expressed ORFs were clustered to detect time-dependent variations in expression and to illustrate expression dynamics after osmotic up-shock. The 15,093 different ORFs showed time-dependent expression patterns (Fig. 32).
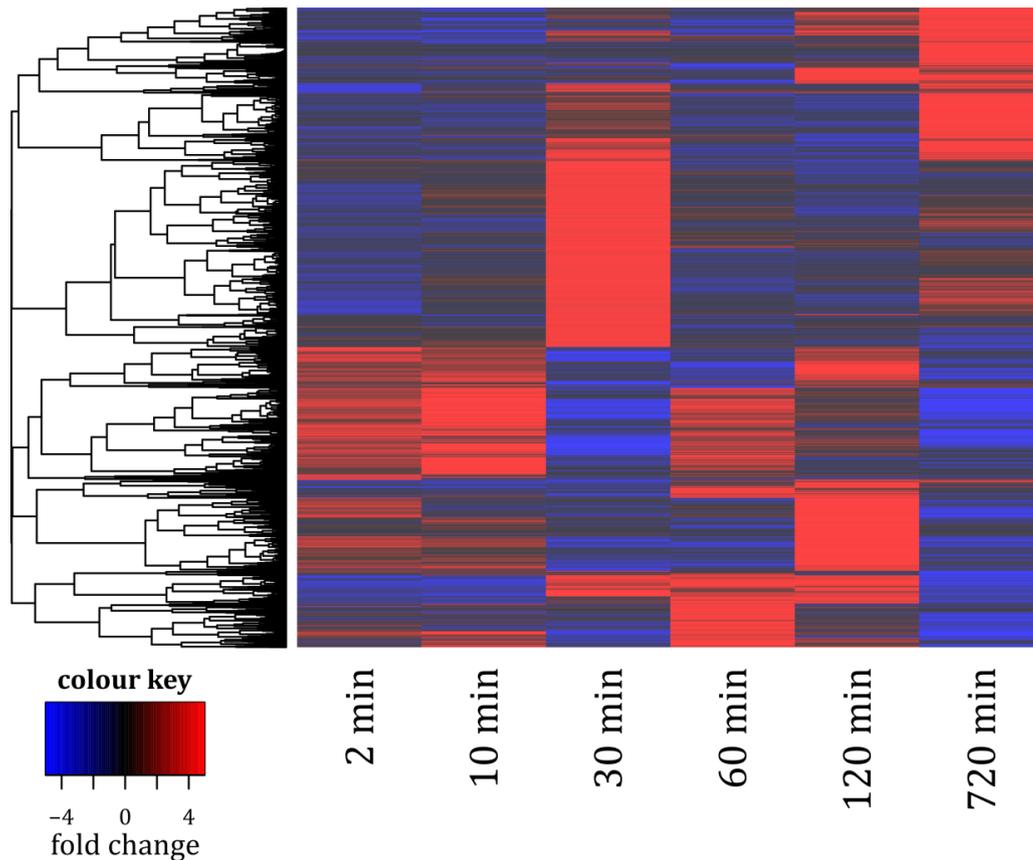
Figure 32: Gene correlation heatmap according to the sampled time points (x-axis) based on the differential expression values of 15,093 ORFs having at least a 4-fold-change expression at any investigated time point compared to control group (t0). Each line represents an individual ORF. The colour-code indicates up- or down-regulation demonstrating a time-dependent pattern of the intracellular expression of *S. salinarum*. The similarity of expression compared to t0 of each ORF is illustrated on the left-hand side by a distance tree. Supplemental data are given in the digital appendix in file 'Input_data_Fig.32_regulation_15093_diffexpr_ORFs'.

To investigate potential metabolic processes, which could be assigned to the above observed time-dependent expression pattern, the 3,202 up- or down-regulated ORFs, which could be annotated to the KEGG database, were ordered according to their KEGG categorization (Fig. 33). Considering the quantity of up- or down-regulated ORFs at a certain time point in dependence of the different categories, the most abundant category was 'genetic information processing' containing 21 – 46 % of all up-regulated and 20 – 42 % of all down-regulated ORFs according to the time point. Comparing now the proportion of ORFs, which were up- or down-regulated at a certain time point, within a specific category with each other, time-dependent differences could be observed (Fig. 33). For the category 'unclassified', the largest proportion of up-regulated ORFs was found at 60 min after salt up-shock followed by 720 min and the largest proportion of down-regulated ORFs was observed after 10 min. For 'organismal

systems', the largest amount of up-regulated ORFs were detected after 30 min followed by 2 min (Fig. 33). The proportion of down-regulated ORFs within the category 'organismal system' fluctuated not significantly (1 – 2 %) but no down-regulated ORFs were ordered in this category after 120 min. Within the category 'cellular processes', the relative proportion of up-regulated ORFs decreased markedly to 4 % at 60 min and the fewest relative proportion of down-regulated ORFs was observed after 2 min lacking any down-regulated ORF after 120 min, which could be placed to this category (Fig. 33). In the category 'environmental processing' the greatest relative proportion of down-regulated ORFs was detected after 2 min (14 %) and the largest relative amount of up-regulated ORFs was observed after 720 min (13 %). Within the category 'genetic information processing', the largest relative proportion of up-regulated ORFs was found at 2 min (46 %) followed by 120 min (40 %) and the largest proportion of down-regulated ORFs was observed after 30 (42 %) and 60 min (37 %). For the category 'biosynthesis of other secondary metabolites', only three up-regulated and one down-regulated ORF could be found (Fig. 33). Also a low number of up- (n = 34) and down-regulated ORFs (n = 5) could be placed to the category 'metabolism of terpenoids and polyketides' (Fig. 33). Within the category 'metabolism of cofactors and vitamins', the relative proportion of up-regulated ORFs were almost constant (3 – 4 %) except the missing of any ORF at 2 min. For down-regulated ORFs placed to the same category, the largest relative proportion was found at 720 and 10 min (6 %) after salt up-shock (Fig. 33). In total, only 13 down-regulated and 43 up-regulated ORFs out of 3,202 ORFs were ordered to the category 'glycan biosynthesis and metabolism' (Fig. 33). The relative proportion of up-regulated ORFs within the category 'metabolism of other amino acids' fluctuated time-dependently between 1 – 2 % showing no significant difference between the certain time points. But the largest relative proportion of down-regulated ORFs within the category 'metabolism of other amino acids' was found at 2 min (4 %). The largest relative proportion of down-regulated ORFs belonging to 'amino acid metabolism' was detected after 2 and 120 min (9 and 15 %). The relative amount of up-regulated ORFs of the category 'amino acid metabolism' varied not clearly (4 – 7 %; Fig. 33). Within the category 'nucleotide metabolism', no down-regulated ORF could be detected at 120 min and the largest reative proportion of up-regulated ORFs was observed after 30 min (6 %). For the 'lipid metabolism', the relative proportion of down-regulated ORFs was almost constant fluctuating between 3 and 5 % (Fig. 33). But the largest relative proportion of up-regulated ORFs belonging to the 'lipid metabolism' was

found at 10 and 30 min (5 and 6 %). Within the category 'energy metabolism', the relative proportion of up-regulated ORFs varied not markedly between the time points (4 – 6 %) but the largest relative proportion of down-regulated ORFs was detected at 120 min (21 %). For the category 'carbohydrate metabolism', the largest proportion of up-regulated ORFs was found after 30 min (10 %) and the largest proportion of down-regulated ORFs after 10 min (12 %; Fig. 33). Summarising the most distinctive differences between the six time points within the certain categories, the down-regulation of the energy and amino acid metabolism after 120 min as well as the up-regulation of 'genetic information processing' after 2 min and the down-regulation after 30 min were the most remarkable comparisons (Fig. 33).
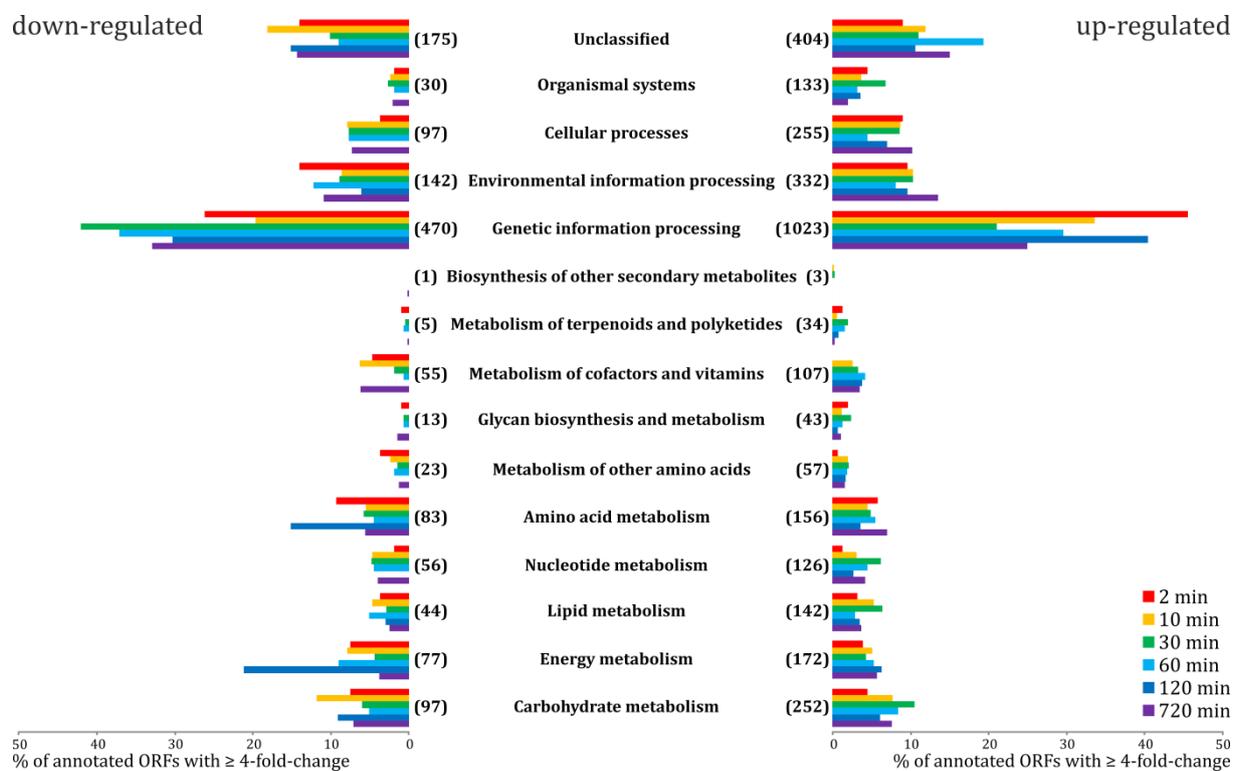
Figure 33: Relative distribution of up- or down-regulated ORFs with at least 4-fold-change expression compared to the control group (t0), which were annotated and ordered to KEGG functional categories. Numbers of up- or down-regulated ORFs per category are written in brackets. The different colours symbolise the respective time point after salt up-shock. Supplemental data are given in the digital appendix in file 'Input_data_Fig.33_KEGG'.

To reveal compartments, where the cellular response to the salt-up shock was most active, the differentially expressed 15,093 ORFs were grouped according to their sub-cellular localisation. Mitochondria, the cytoplasm and the nucleus appeared as the cell compartments, in which the most proteins were up- or down-regulated during the

720 min experiment (Fig. 34). When the different cell compartments were investigated according to the six different time points after salt up-shock, the most up-regulation of ORFs in the nucleus was found after 2 and 10 min (Fig. 34). The significant greatest amount of up-regulated ORFs was detected in the mitochondrion after 30 min followed by a dominantly increase of down-regulated ORFs equally in the mitochondrion after 120 min (Fig. 34). In between, the highest amount of cytoplasmic ORFs was up-regulated at 60 min after salt up-shock and the most cytoplasmic ORFs with a down-regulation were detected after 720 min (Fig. 34). For the cell membrane as locality with direct contact to the environment, the largest proportion of up-regulated ORFs was found at 2 min, 10 min and 120 min in contrast to 30 min, 60 min and 720 min (Fig. 34).
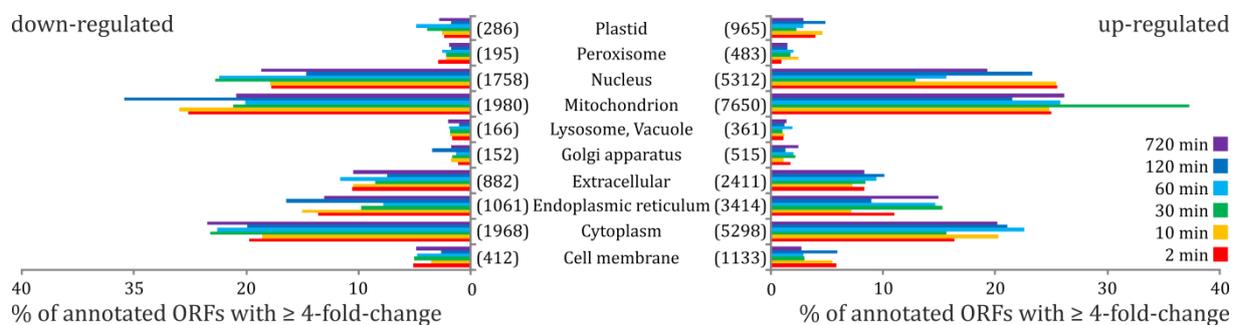
Figure 34: Relative distribution of up- or down-regulated ORFs with at least 4-fold-change expression compared to control (t0) ordered by their predicted sub-cellular location. Numbers of up- or down-regulated ORFs per predicted sub-cellular location are written in brackets. The different colours illustrate the six different time points after salt up-shock. Supplemental data are given in the digital appendix in file 'Input_data_Fig.34_deeploc'.

For in-depth investigations targeting ORFs specifically relevant for haloadaptation strategies, differentially expressed ORFs were sorted and if possible, assigned to one of the following categories: 'sensing', 'compatible solute transport', 'compatible solutes' and 'ion transport'. ORFs, which showed exactly the same functional application within a category, were combined by the arithmetic mean of their time-dependent expression pattern. The following heatmaps in the next sub-chapters illustrate specific intracellular reactions of *S. salinarum* in a time-dependent resolution to elucidate intracellular dynamics after salt up-shock until cell acclimatisation.

3.2.1 Sensing mechanisms in *S. salinarum*

Several components for sensing and cell signalling were found in the transcriptome of *S. salinarum* and their expression differed over time. First, an oxygen sensor histidine kinase and a histidine kinase of a two-component sensing system were strongly up-regulated after 2 and 10 min (Fig. 35). Likewise, other histidine kinases for signal transduction and hybrid sensors were also up-regulated at 10 min together with cAMP-regulated phosphoproteins as well as cAMP-activated global transcriptional regulators (Fig. 35). After 30 min, most dominant up-regulation was detected for cyclic di-AMP synthases. Additionally, the up-regulation for mitogen-activated protein kinase phosphatase, inositol-tetrakisphosphate 1-kinase, glutamate receptor as well as cAMP-regulated D2 protein was less strong but significant (Fig. 35). The expression values of two-component system sensor histidine kinase, thymidine kinase, signal transduction histidine kinase, oxygen sensor histidine kinase and cAMP-activated global transcription regulator after 30 min appear conspicuous compared to the respective values at 10 and 60 min (Fig. 35). The biological relevance is discussed later and could be tested by further investigations using quantitative real-time PCR (qRT-PCR). After 60 min, additional regulatory proteins of the two-component system were strongly up-regulated together with the sensor histidine kinase of the two-component system, an oxygen sensor histidine kinase and cAMP-activated global transcriptional regulators (Fig. 35). Thioredoxin and thioredoxin reductases were significantly up-regulated at 120 min after salt up-shock. At the last investigated time point (720 min), osmotically inducible proteins were most up-regulated followed by response regulators and their receiver proteins (Fig. 35). Also up-regulated were thymidine kinases and thioredoxin-disulfide reductases. In contrast, a strong down-regulation could be detected for sensor histidine kinases of the two-component system, signal transduction histidine kinases, cAMP-dependent protein kinases and cAMP-activated global transcriptional regulators (Fig. 35).
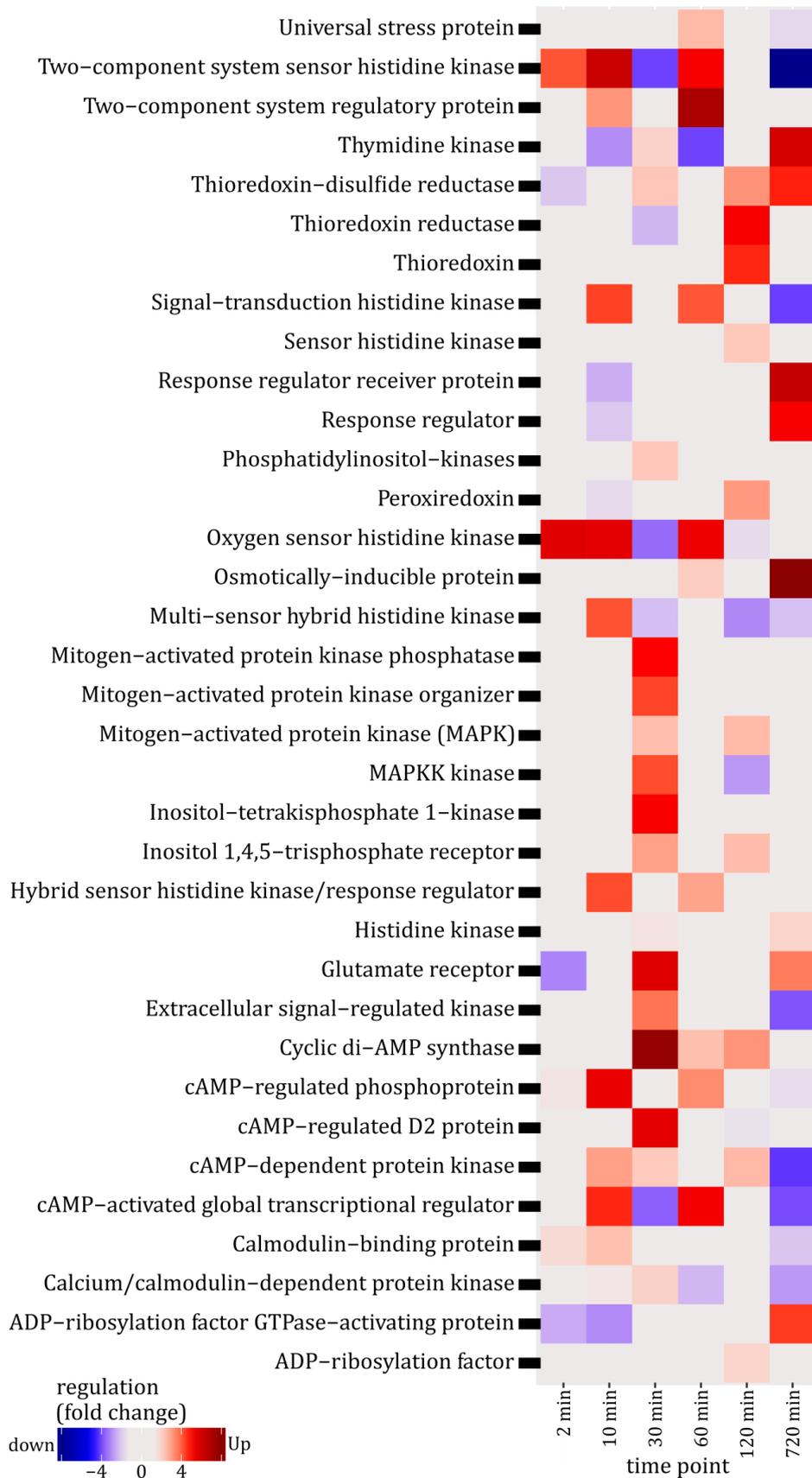
Figure 35: Expression heatmap of the different proteins with their annotated functions (e-value threshold 1e^{-5}) belonging to the main category 'sensing'. Several components for sensing and cell signalling were found in the transcriptome of *S. salinarum* and their expression differed over time. Supplemental data are given in the digital appendix in files 'Annotationtable_15093_diffexpr_ORFs' (protein identifiers) and 'Input_data_Fig.35_sensing' (expression values).

3.2.2 Compatible solute transport in *S. salinarum*

Nine different putative transporters for potential compatible solutes were detected in the transcriptome of *S. salinarum* and differentially expressed showing a time-dependent expression pattern (Fig. 36). Using the machine learning algorithm of Deeploc-1.0 (Almagro Armenteros et al., 2017) for prediction of the sub-cellular localisation of annotated proteins, five transporters were likely located in the cell membrane and three in the lysosomes or vacuoles. Additionally, one choline transporter was predicted to be located in the mitochondrion (Fig. 36). In an early reaction, the capacity for choline transport over the cell membrane into the cytoplasm increased at 2 and 10 min (Fig. 36). Simultaneously, the transport capacity for sucrose over the cell membrane and UDP-N-acetylglucosamine into the lysosomes or vacuoles decreased slightly at 2 min. After 30 min, the capacity of Na^+/*myo*-inositol co-transporter and sucrose transport over the cell membrane as well as the capacity for the import of UDP-N-acetylglucosamine into the lysosomes or vacuoles strongly increased (Fig. 36). A very strong up-regulation of the transport system for glycine betaine located at the lysosomes or vacuoles was observed after 60 min as well as after 120 min. A slightly increased capacity for alanine/cation symport over the cell membrane was also detected after 60 min (Fig. 36). After 120 min, a very distinct increase in the transport capacity for choline into the mitochondrion was measured (Fig. 36). At 720 min after salt up-shock, the capacity for proline transport into the cytoplasm as well as for the osmoregulated proline transport into lysosomes or vacuoles increased significantly. In contrast, an obvious down-regulation of Na^+/*myo*-inositol co-transporter, alanine/cation symporter and choline transporter in the cell membrane was also observed after 720 min (Fig. 36).
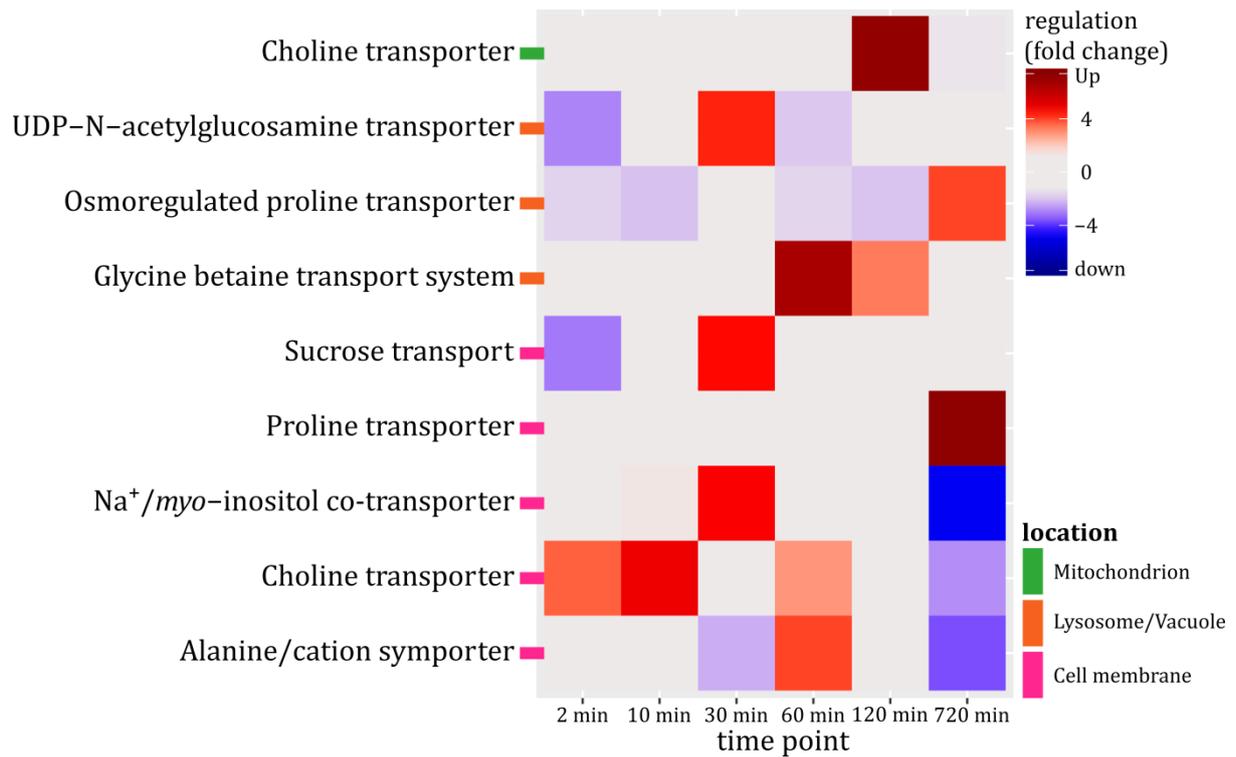
Figure 36: Expression heatmap of different putative transport mechanisms for potential compatible solutes found in the transcriptome of *S. salinarum* (e-value threshold 1e^{-5}). The transport capacity (concluded from corresponding expression values) for the compatible solutes alanine, choline, myo-inositol, proline, sucrose, glycine betaine and N-acetylglucosamine varied over time and was partly location specific. Supplemental data are given in the digital appendix in files 'Annotationtable_15093_diffexpr_ORFs' (protein identifiers) and 'Input_data_Fig.36+38_transport' (expression values).

3.2.3 Compatible solutes and their pathways in *S. salinarum*

Former experiments of Weinisch et al. (2018b) revealed that *S. salinarum* accumulates glycine betaine (GB) and ectoine (Ect) for osmotic adjustment. The experiments also indicated that *S. salinarum* is able to synthesize or import GB and Ect but the required enzymes could not be identified so far. Therefore, the transcriptome of *S. salinarum* was searched for enzymes of GB and Ect synthesis and other potential compatible solutes, which were known from literature. In the transcriptome of *S. salinarum*, enzymes of pathways for the synthesis or catabolism of seven putatively used compatible solutes were found showing time-dependent expression patterns. The potential compatible solutes were proline, *myo*-inositol, glycerol, glutamate, glycine betaine (GB), ectoine (Ect) and choline. Interestingly, the complete functional pathway for the *de novo* synthesis of ectoine (Ect) was detected in *S. salinarum* starting with aspartate followed by the usage of the five different enzymes aspartate kinase (AsK), aspartate-

semialdehyde dehydrogenase (AsD), diaminobutyrate-2-oxoglutarate aminotransferase (EctB), diaminobutyrate acetytransferase (EctA) and ectoine synthase (EctC; Suppl. Fig. 8). But only one specific enzyme for ectoine synthesis was differentially expressed over time. The enzyme diaminobutyrate acetytransferase (EctA) of the ectoine synthesis pathway was first slightly down-regulated after 30 min and then significantly up-regulated at 60 and 120 min after salt up-shock (Fig. 37).

For choline, two synthesis enzymes were up-regulated at 10 and 30 min, but in parallel choline dehydrogenase were also up-regulated after 30 min symbolising the first step of glycine betaine (GB) synthesis from the precursor choline (Fig. 37). The second enzyme for the synthesis of the compatible solute glycine betaine, betaine-aldehyde dehydrogenase (BADH), was up-regulated only after 60 min (Fig. 37). Hence, the complete pathway for the two-stage *de novo* synthesis of glycine betaine (GB) was found in the transcriptome of *S. salinarum* containing both enzymes choline dehydrogenase (CDH) and betaine-aldehyde dehydrogenase (BADH; Suppl. Fig. 6).

Considering the other putatively used compatible solutes, enzymes for glutamate synthesis as well as glutamate catabolism were up-regulated after 10 min until 720 min. At 720 min, enzymes for both processes were down-regulated (Fig. 37). In contrast, the differentially expressed enzyme monoglyceride lipase for glycerol synthesis was exclusively down-regulated after 30 and 720 min. Enzymes for glycerol catabolism were up-regulated simultaneously at the same time points (Fig. 37). A time-dependent shift in the regulation of synthesis and catabolism was found for *myo*-inositol. *Myo*-inositol-2-dehydrogenase as part of the catabolic pathway was up-regulated after 30 min and the opposing enzyme inositol-3-phosphate synthase for the synthesis of *myo*-inositol was up-regulated after 120 min (Fig. 37). Enyzmes for proline synthesis were slightly down-regulated after 10 and 30 min and a strong and increasing up-regulation could be measured after 60 min until inclusively 720 min (Fig. 37).
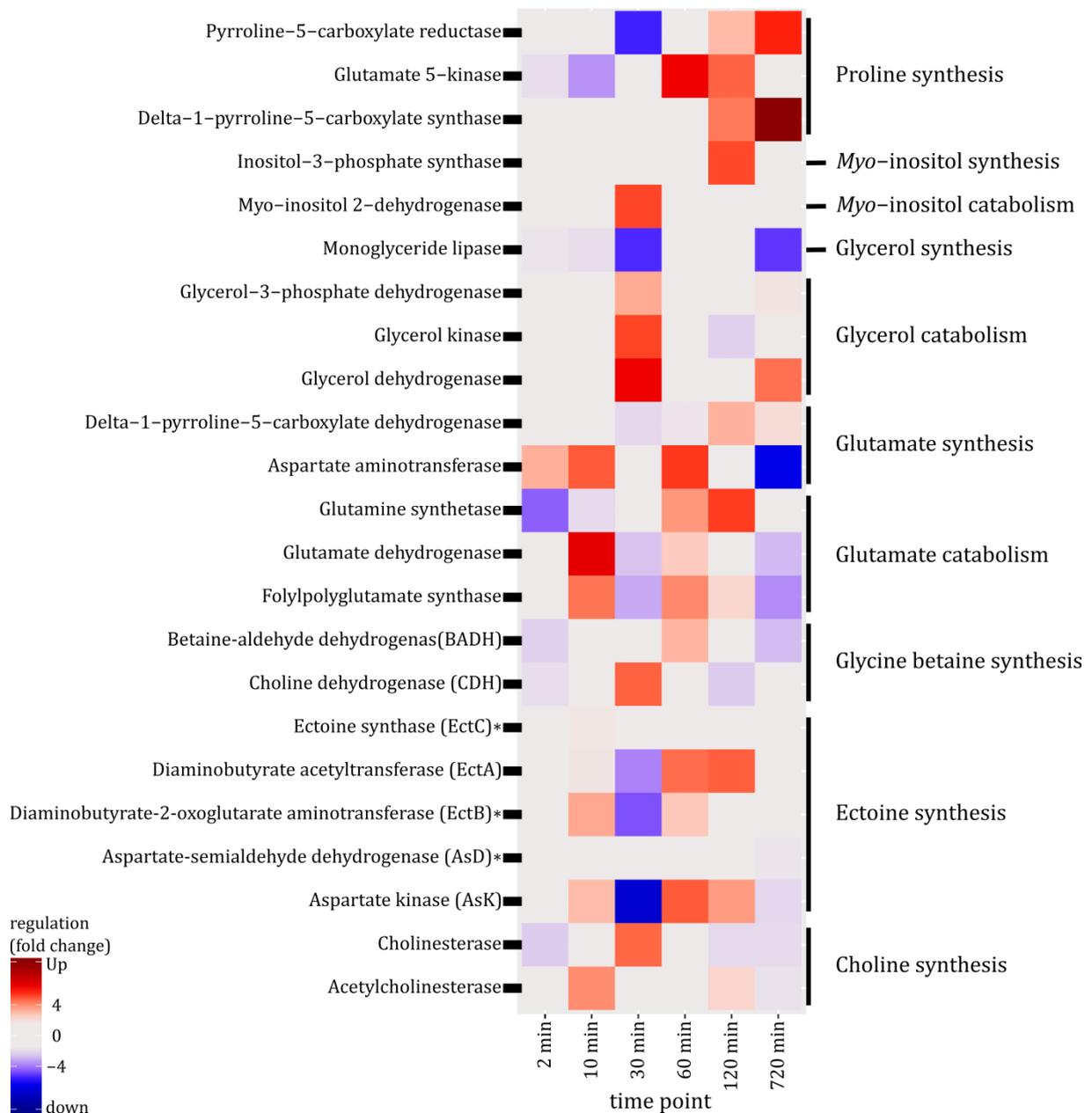
Figure 37: Expression heatmap of the different enzymes for synthesis or catabolism of the potential compatible solutes choline, ectoine, glycine betaine, glutamate, glycerol, *myo*-inositol and proline found in the transcriptome of *S. salinarum* (e-value threshold $1e^{-5}$). According to the different time points after salt up-shock, different enzymes were up- or down-regulated. Enzymes showing no significant expression differences after salt up-shock at the six investigated time points were marked (*), namely AsD, EctB and EctC. Supplemental data are given in the digital appendix in files 'Annotationtable_15093_diffexpr_ORFs' and 'Input_data_Fig.37_compatible_solute_synthesis' (protein identifiers and expression values).

3.2.4 Ion transport in *S. salinarum*

Thirteen different putative ion transporters and transport mechanisms were detected in the transcriptome of *S. salinarum* demonstrating nineteen possibly different ion transport capacities embedded at different predicted cell components with a time-

dependent expression pattern. As an early response to a salt up-shock, mechanosensitive ion channels located in the plastid were strongly up-regulated after 2 and 10 min. After 30 min, ion channels of the endoplasmic reticulum, namely voltage-gated K^+ channel, mechanosensitive ion channel, K^+ channel, Cl^- channel, Ca^{2+} channel and anion channel, were significantly up-regulated (Fig. 38). A similar up-regulation was observed for $Na^+/P_{inorganic}$ co-transporter, mechanosensitive ion channel and Cl^-/fluoride channel located in the lysosomes or vacuoles after 30 min. In contrast, Na^+/K^+ transporting ATPases and K^+ transporting ATPases of the cell membrane were significantly down-regulated also after 30 min (Fig. 38). After that, an increasing capacity for Na^+ expulsion was detected by the strong up-regulation of Na^+/H^+ antiporters located in the cell membrane at 60 and 120 min. This Na^+ expulsion capacity after 120 min was complemented by the up-regulation of mechanosensitive ion channels in the cell membrane as well as K^+ channels and Cl^- channels located in the endoplasmic reticulum (Fig. 38). After 720 min, mitochondrial Ca^{2+} transporting ATPases, Na^+/K^+ transporting ATPases and Na^+/H^+ antiporters of the endoplasmic reticulum as well as cation transporting ATPases of the cell membrane were strongly up-regulated. Simultaneously, Ca^{2+} channels in the cell membrane and K^+ channels in the endoplasmic reticulum were significantly down-regulated (Fig. 38).
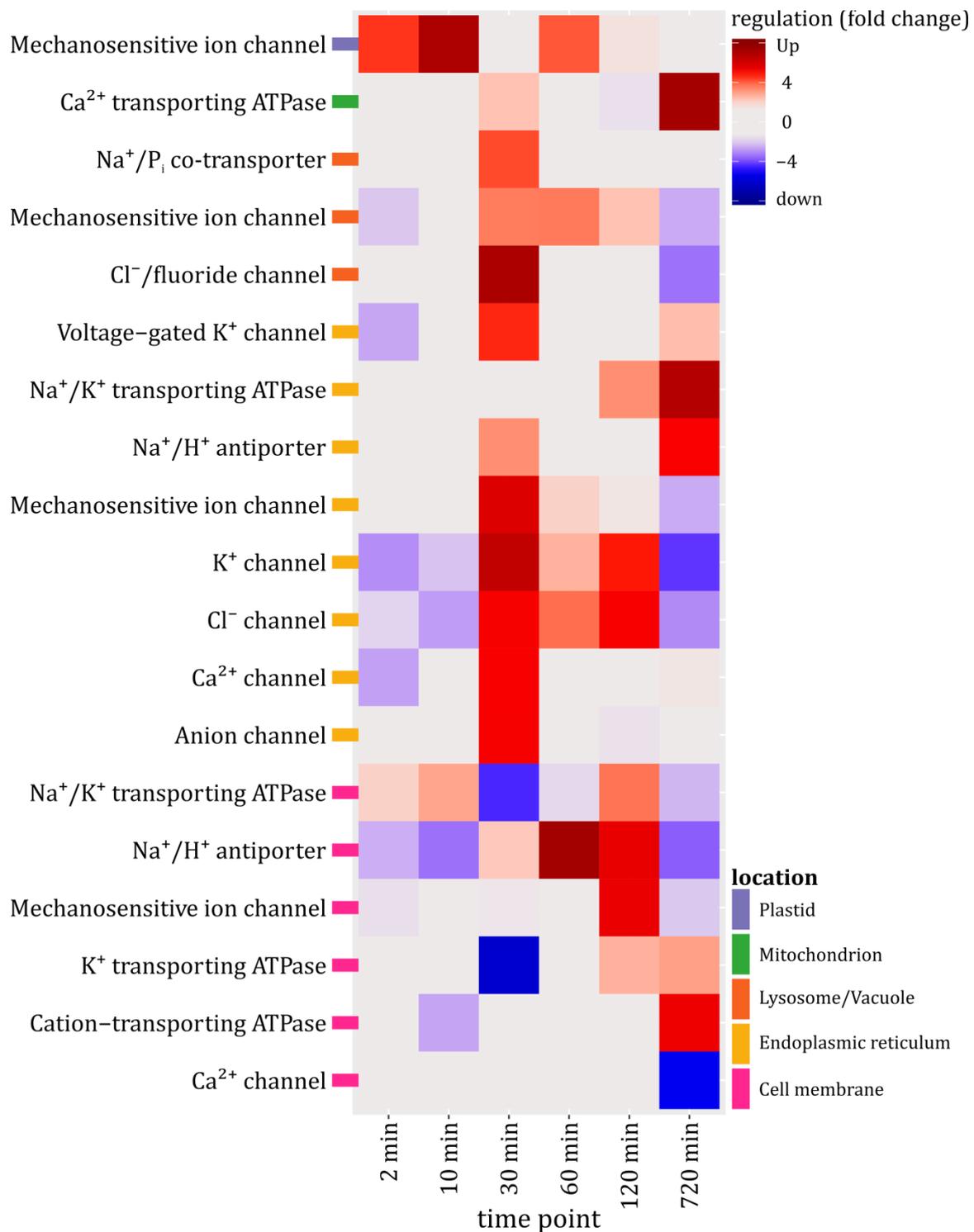
Figure 38: Expression heatmap of the different putative ion transport mechanisms found in the transcriptome of *S. salinarum* (e-value threshold 1e^{-5}). The ion transport capacities (concluded from corresponding expression values) varied over time and were mostly location dependent. Six transport mechanisms were possibly located in the cell membrane, eight in the endoplasmic reticulum, three in the lysosome or vacuole and respectively one in the mitochondrion and plastid. Supplemental data are given in the digital appendix in files 'Annotationtable_15093_diffexpr_ORFs' (protein identifiers) and 'Input_data_Fig.36+38_transport' (expression values).

3.2.5 Other processes potentially linked to haloadaptation in *S. salinarum*

Besides the former investigated transport mechanisms for compatible solutes or ions, the sensing machineries, intracellular signal transduction cascades and the synthesis pathways of compatible solutes, additional possibly relevant intracellular processes and metabolic functions regarding haloadaptation in *S. salinarum* were found (Fig. 39). These included ubiquinone synthesis, tricarboxylic acid (TCA) cycle, sterol biosynthesis, reactive oxygen species (ROS) detoxification, protein damage control, phosphatidyl-ethanolamine synthesis, O_2 respiration, heme synthesis, fermentation and cyclic nucleotide-dependent signalling (Fig. 39). Enzymes of the ubiquinone synthesis were slightly down-regulated at 2 min and almost not differentially expressed after 10 min, but again strongly up-regulated between 30 and 120 min after salt shock. At 720 min, the capacity of ubiquinone synthesis decreased again (Fig. 39). Another heterogenous pattern was found for the TCA cycle. The pyruvate dehydrogenase kinase and fumarate hydratase were down-regulated first after 2 min and then most dehydrogenases were up-regulated between 10 and 720 min (Fig. 39). One exception was 2-methylisocitrate lyase, which was constantly up-regulated at all investigated time points (Fig. 39). In contrast, a higher capacity for sterol biosynthesis was only found with the enzyme sterol-24-C-methyltransferase after 120 min (Fig. 39). Concluding the expression pattern of the enzymes for ROS detoxification, the process of ROS detoxification had only marginal increased capacity after 120 and 720 min (Fig. 39). However, the protein damage control was targeted by several different heat shock proteins in *S. salinarum* and demonstrated different changes of up- or down-regulation for different heat shock proteins (Fig. 39). A predominant up-regulation of the enzymes for the heme synthesis was observed as well as for the process of O_2 respiration (Fig. 39). Equally predominantly up-regulated were the enzymes for fermentation between 10 and 60 min after salt up-shock as well as the different phosphodiesterases belonging to cyclic nucleotide-dependent signalling, which were significantly up-regulated at all time points except at 720 min (Fig. 39). Furthermore, the phosphoethanolamine-cytidyltransferase of the phosphatidyl-ethanolamine synthesis as the last additional group was up-regulated only after 30 min (Fig. 39).
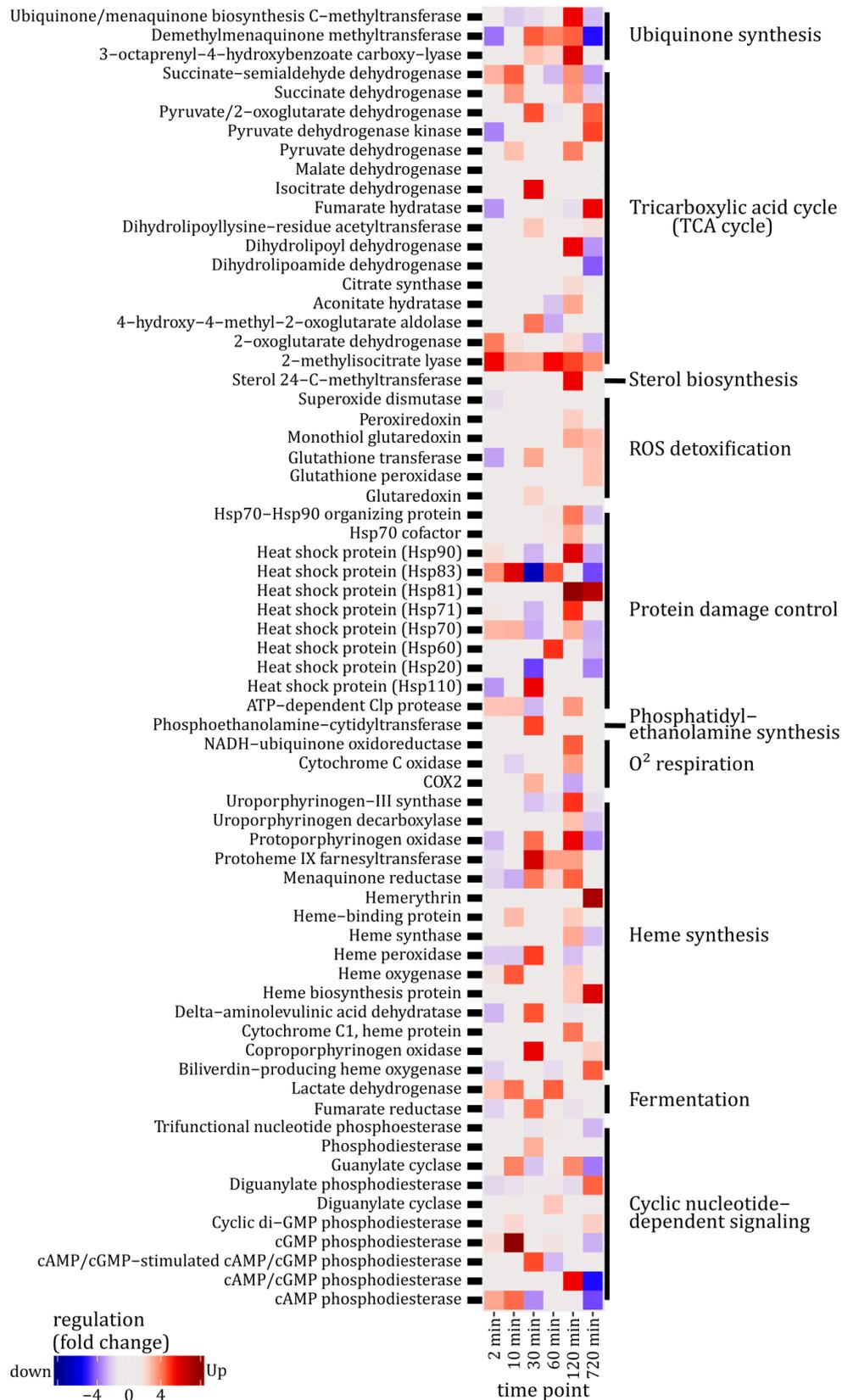
Figure 39: Expression heatmap of possibly relevant intracellular processes and metabolic functions regarding haloadaptation in *S. salinarum* (e-value threshold 1e^{-5}) sorted into the functional groups ubiquinone synthesis, tricarboxylic acid (TCA) cycle, sterol biosynthesis, ROS (reactive oxygen species) detoxification, protein damage control, phosphatidyl-ethanolamine synthesis, O$_2$ respiration, heme synthesis, fermentation and cyclic nucleotide-dependent signalling. Supplemental data are given in the digital appendix in files 'Annotationtable_15093_diffexpr_ORFs' (protein identifiers) and 'Input_data_Fig.39_other_processes' (expression values).

3.3   Comparing the acidity of proteomes and cell compartments

The isoelectric points (pI) of the predicted proteins of *S. salinarum* were compared to isoelectric point profiles of proteins from the halophile bacterium *Salinibacter ruber*, the halophile black yeast *Hortaea wernickii* and the freshwater ciliate *Oxytricha trifallax* to detect differences between the proteome acidity of these three organisms and *S. salinarum*. Because these three organisms use different haloadaptation strategies, it can be assumed if the transcriptome of *S. salinarum* reveals a similar proteome acidity pattern, *S. salinarum* could be able to use the same haloadaptation strategy.

The proteome of *O. trifallax*, being a freshwater representative using no haloadaptation strategy, consisted of proteins with mostly alkaline isoelectric points (47.6 % of proteins above pI 8), with a peak at pI 9.6 (11.5 %), and a rather minor fraction of proteins with more acidic values (27.5 % of proteins below pI 6; Fig. 40). In contrast, the isoelectric point profile of the proteome of *H. wernickii* appeared to be shifted towards a higher acidity as most of its proteins (52.0 %) displayed acidic pI values ($\leq$ 6), peaking at a pI of 6 (11.2 %; Fig. 40). Over 74 % of all proteins from *S. ruber* revealed high acidic isoelectric point values ($\leq$ 6), with most proteins (24.7 %) having a pI of 4.4 (Fig. 40). Interestingly, the majority of proteins (69.4 %) from *S. salinarum* had more alkaline isoelectric points ($\geq$ 8), with most proteins (9.4 %) having a pI of 10. Logically, less than 18.9 % of *S. salinarum´s* proteins were connected to acidic values ($\leq$ 6; Fig. 40). Investigating only the 15,093 ORFs of *S. salinarum* having a 4-fold-change at least once compared to t0, most proteins had alkaline isoelectric points (60.1 % of proteins $\geq$ pI 8) showing a double peak at pI 9.6 and 10 (7.8 % each; Fig. 40). The only difference of the subset of differently regulated ORFs compared to the complete transcriptome of *S. salinarum* was observed for the proteins with more acidic values ($\leq$ 6). Their amount increased from former 18.9 % to 27.4 % (Fig. 40).
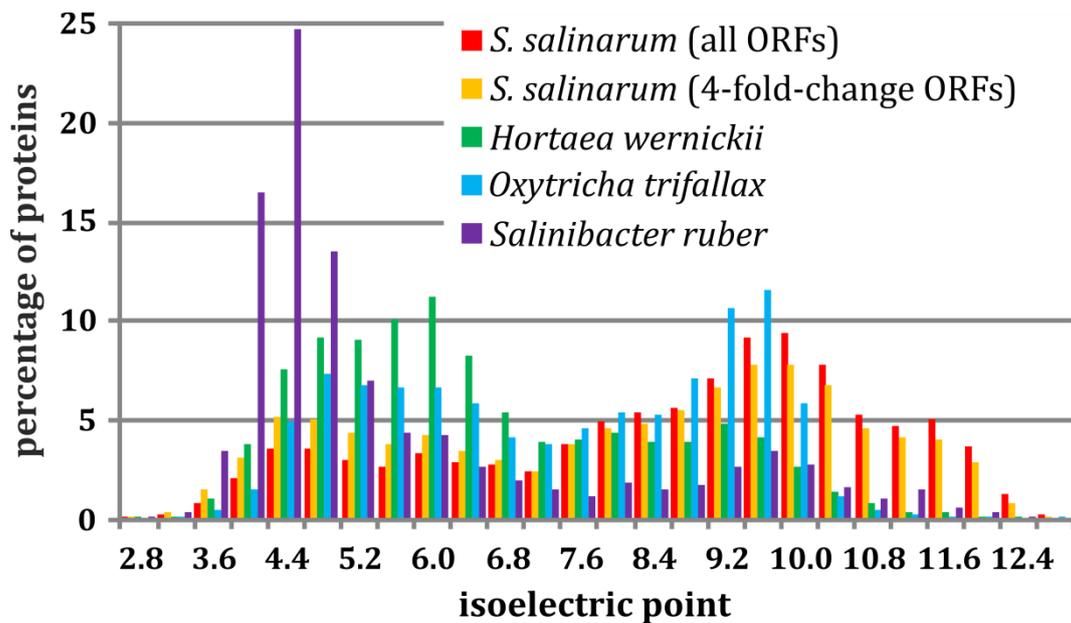
Figure 40: Isoelectric points of the proteins from *Schmidingerothrix salinarum, Salinibacter ruber*, *Hortaea wernickii* and *Oxytricha trifallax*. The proteins of *S. salinarum* originated from this study, the proteins from the other organisms derived from entries of NCBI´s non-redundant protein reference database (NCBI- RefSeq database; http://www.ncbi.nlm.nih.gov/RefSeq/; O'Leary et al., 2016). *S. ruber* is a 'salt-in' strategist, *O. trifallax* lives in freshwater environments and *H. wernickii* has a lot of fungal specific haloadaptation mechanisms.

Considering the 168,454 proteins of the transcriptome of *S. salinarum* were located in different cell compartments, the distribution of the isoelectric points (pI) of the predicted proteins were analysed according to their sub-cellular localisation. Initially, the relative amount of proteins sharing the same pI value varied markedly between cell compartments (Fig. 41). The lysosomes and vacuoles contained the largest proportion of proteins showing strong acidic values (64.1 % of proteins ≤ pI 6), with a peak at pI 4.8 (18.6 %). In contrast, the majority of proteins located in the mitochondrion (91.3 %) as well as in the plastid (81.7 %) exhibited very alkaline isoelectric points (pI value ≥ 8). Their peaks were observed at pI 10.4 with 14.3 % (mitochondrion) respectively 9.5 % (plastid). Similarly, most of the proteins originated from the peroxisome showed also more alkaline pI values (71.5 % of proteins ≥ pI 8), with a peak at pI 10.0 (12.1 %; Fig. 41). Additionally, the majority of proteins located in the endoplasmic reticulum (61.8 %) and Golgi apparatus (54.5 %) were only slightly alkaline (8 ≤ pI ≤ 10.4), showing peaks at pI 9.6 (11.2 %; endoplasmic reticulum) respectively at pI 10.0 (9.1 %; Golgi apparatus). The majority of proteins originated from the nucleus contained also alkaline pI values (77.6 % of proteins ≥ pI 8), showing a peak at pI 10.0 (10.1 %; Fig. 41). In contrast, the proteins located in the cytoplasm

demonstrated clear peaks (8.8 % and 7.6 %) at acidic pI values (pI 5.2 and pI 5.6) but largest relative proportion of cytoplasmic proteins exhibited alkaline isoelectric points (44.5 % of proteins ≥ pI 8). Only 34.9 % of all cytoplasmic proteins had more acidic pI values (≤ 6; Fig. 41). Actually, the cytoplasm of *S. salinarum* could only be slightly acidic, if at all, and lysosomes as well vacuoles were strongly acidic according to the isoelectric points of their proteins.
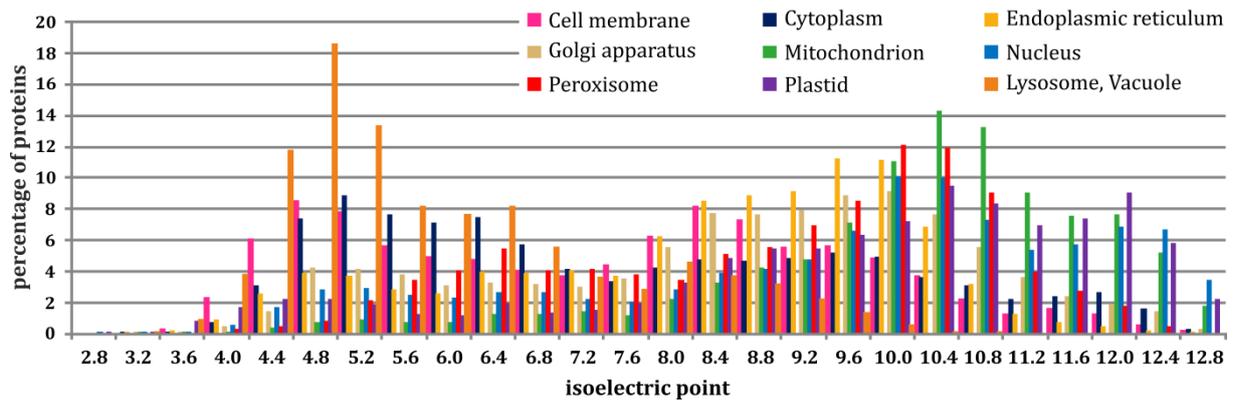
Figure 41: Distribution of the amount of proteins according to their isoelectric point in dependence to their predicted sub-cellular localisation inside *Schmidingerothrix salinarum*. Lysosomes and vacuoles were the cell compartments containing the most acidic proteomes. In contrast, mitochondria and plastids consisted of proteins with the most alkaline isoelectric points. The predicted proteins of the cytoplasm *S. salinarum* provided relatively more acidic values than it was shown before considering all proteins (see Fig. 40).

IV.   Discussion

1.   eDNA metabarcoding of protistan plankton communities in the Baltic Sea

Until now, the existence of salinity boundaries and protistan community shifts were unclear for the transition zone from freshwater to marine waters (salinity 0 – 40 ‰). Although Telesh and colleagues (2011; 2013) predicted a protistan species maximum in the horohalinicum (5 – 8 ‰ salinity) of the Baltic Sea, nothing was known about protistan community structures and responsible environmental factors or how protistan plankton from brackish waters differ from the protistan plankton community detected in hypersaline environments (Filker et al., 2017). Because the study of Telesh et al. (2011) was only a literature survey based on combining microscopic data from different studies, I used high-throughput eDNA metabarcoding of the V4 region of the 18S rDNA in order to get a complete picture of community compositions and protistan diversity.

Investigations of protistan plankton diversity in the Baltic Sea with traditional microscopy methods revealed 2075 recorded phytoplankton species (Hällfors, 2004) and 768 described ciliate species (Telesh et al., 2008; Filker et al., 2019). Although no species counts are available for heterotrophic nano- and picoflagellates occurring in the Baltic Sea because their small sizes hardly allow detailed investigations of their taxonomic features (Filker et al., 2019), the protistan diversity in the Baltic Sea seems to be still very high and is hardly matched by any other water body reported in the published literature (Filker et al., 2019). The most likely explanation for this vast diversity of unicellular planktonic eukaryotes is the unique location of the Baltic Sea. Due to the only connection to the North Sea through Skagerrak and Kattegat, saline seawater enters the Baltic Sea from the west and is mixed with more brackish waters from the east, originating from freshwater input by rivers and streams of the surrounding countries. Therefore a stable west to east salinity gradient is established in the Baltic Sea. Because of this pronounced salinity gradient, giving rise to freshwater, brackish water and marine communities, the huge protistan diversity can be explained emphasizing the relevance of the Baltic Sea as a biodiversity reservoir (Filker et al., 2019).

To investigate community compositions and protistan diversity, samples were taken within a 2-week period along a west to east transect in the Baltic Sea covering a pronounced salinity gradient ranging from 17.3 to 22.1 ‰ in the west to 7.9 to 8.4 ‰ in the east (Fig. 9; Suppl. Tab. 1). Additional freshwater samples were excluded because

such samples would either have to be collected in estuaries and lakes or rivers (Filker et al., 2019). Estuaries, however, are subjected to a different dynamic with quickly changing environmental conditions, variable freshwater input and short hydrological retention with limited possibilities for the establishment of an autochthonous stable plankton community (Crump et al., 2004). For rivers and lakes instead, the diversity of plankton communities from these discrete freshwater bodies varies greatly (Grossmann et al., 2016; Boenigk et al., 2018) and therefore are difficult to compare with samples from a continuous salinity gradient (Filker et al., 2019).

Despite ignoring freshwater protistan plankton communities, clear patterns along the marine – brackish water salinity gradient were observed. Alpha-diversity increased significantly with decreasing salinity and reached its maximum in the brackish water zone, which could be demonstrated with a significant correlation between the Simpson index, the effective number of species and the measured salinity (Fig. 10; Tab. 4; Filker et al., 2019). The same species maximum in brackish waters might apply to bacteria because Pavloudi et al. (2017) reported recently about high species numbers for brackish bacteria in coastal waters of the Ionian Sea. For that reason, it could be suggested that microorganisms follow fundamentally different rules than multicellular organisms when it comes to distribution patterns along salinity gradients (Filker et al., 2019). Additionally, the origin of protistan species living in salinities below 10 ‰ could be explained by two scenarios. First, locally adapted freshwater or marine populations were predominant in brackish water while only a few truly brackish species occurred (Filker et al., 2019). Second, the protistan plankton was dominated by a brackish protistan microbiome, rather than locally adapted freshwater or marine populations resulting in a lower diversity (Filker et al., 2019). In general, the central and eastern parts of the Baltic Sea have more constant living conditions for protistan plankton because natural boundaries like Darss Sill prevent a regular mixing with higher saline waters originating from the western Baltic Sea and the North Sea. Interestingly, bacterioplankton inhabiting the central Baltic Sea comprises members of a global brackish metacommunity, that were adapted to these specific brackish salinities long before the formation of the Baltic Sea and originated not from local freshwater or marine bacterial populations (Hugerth et al., 2015; Filker et al., 2019). To test this possibility for protistan plankton, subsequent fragment recruitment (Niu et al., 2011) is recommended on large-scale metagenome comparisons of protistan plankton from the

Baltic Sea and brackish waters, lakes and oceans from other parts of the world (Filker et al., 2019).

After data processing of the eDNA metabarcoding samples for protistan plankton in the Baltic Sea, the number of OTUs (343 in average) and the number of high-quality target reads per sample were satisfying according to sampling depth indicated by reaching almost the asymptote in the rarefaction analyses (Suppl. Fig. 3). This guaranteed that the received results are reliable and not based on sample biases. The high amount of detected protistan OTUs (n = 2703) in this study, which were assigned to ciliophora (n = 87), dinophyceae (n = 1838), dictyochophyceae (silicoflagellates, n = 505) and bacillariophyceae (diatoms, n = 75), reflected the high diversity within the protistan group (Filker et al., 2019). As comparison, Hu et al. (2016) sequenced planktonic organisms in the Baltic Sea also on a large-scale and detected in total 1860 OTUs including metazoans and fungi. Interestingly, the number of OTUs belonging to ciliophora from Hu et al. (2016) was nearly twice as high as in my study (Filker et al., 2019). But Hu et al. (2016) reported that only 548 OTUs could be assigned to pigmented protistan lineages, predominantly dinophyceae (238 OTUs), chrysophyta (100 OTUs), chlorophyta (99 OTUs) and bacillariophyceae (55 OTUs), whereas in my study the 1838 dinophyceae OTUs alone far exceed this number (Filker et al., 2019).

Although the protistan plankton diversity in the Baltic Sea has been recorded for more than a century, both this study and Hu et al. (2016) as two massive sequencing studies contained a higher diversity compared to the numbers of recorded taxa from microscopy surveys (Hällfors, 2004; Telesh et al., 2008). These discrepancies could result from a combination of technical biases and biological differences, such as seasonal successions and blooms, but the respective contribution of different factors is hard to quantify (Filker et al., 2019). Furthermore, the technical biases of microscopy and molecular protistan diversity studies are well known and have been discussed before by several authors in detail (see e. g. Bachy et al., 2013; Stoeck et al., 2014; Visco et al., 2015; Rivera et al., 2018). Additional examples for discussions of biases including PCR biases, metabarcode regions, taxonomic classifications, OTU cluster algorithms, (relative) abundance patterns, and also the potential influence of ancient DNA in molecularbased diversity studies (Filker et al., 2019) can be found at Stoeck et al. (2010), Pinto and Raskin (2012), Bachy et al. (2013), Stoeck et al. (2014), Hu et al. (2015), Forster et al. (2016), Singer et al. (2016) and will not be repeated here.

Regarding biological differences, the protistan community composition varied clearly between samples and demonstrated a salinity-dependent pattern (Fig. 11) similar to alpha-diversity, which increased with decreasing salinity. In general, dinophyceae, bacillariophyceae and dictyochophyceae belong to the most important phytoplankton classes in the Baltic Sea, sensitively reacting to seasonal changes by forming extensive spring blooms, predominantly in the western Baltic Sea (Wasmund et al., 2013; Wasmund et al., 2017). The typical succession patterns consist of bacillariophyceae in early spring and dinophyceae and dictyochophyceae in late spring (Filker et al., 2019). In the eastern parts instead, dinophyceae outnumber the declining diversity of bacillariophyceae and dictyochophyceae (Fig. 11; Wasmund et al., 2013; Wasmund et al., 2017). Additionally, Wasmund et al. (2013) reported also from the eastern parts of the Baltic Sea that ciliophora co-increase with dinophyceae to compensate for the lower diversity of bacillariophyceae and dictyochophyceae, corrobating well with this metabarcoding study and its sampling in mid-April (Filker et al., 2019). Therefore, it could be possible that seasonal succession was also mirrored in the detected community patterns. But the 2-week sampling period of this study as a relatively short timeframe made this possibility unlikely. Statistically, no significant correlation was found between sampling time and community structures along the spatial sampling gradient (envfit analysis of NMDS; Tab. 5). Furthermore, the described macroecological patterns of different taxon groups in the metabarcoding study of Hu et al. (2016) did also not match the typical spring picture of dominant microplankton in the Baltic Sea due to sampling in mid-summer (Filker et al., 2019). This confirmed the minor importance of potential seasonal succession for the observed community patterns in this study.

According to community composition, the protistan plankton of the Baltic Sea could be divided in dictyochophyceae dominated communities at salinities > 12 ‰ west of the Darss Sill and dinophyceae dominated communities at salinities < 10 ‰ in the east. Herlemann et al. (2011) also identified the same salinity barrier for bacteria in a detailed metabarcoding study of bacterioplankton along a Baltic Sea salinity gradient. Even with a better resolution based on OTUs instead of taxonomic ranks, the beta-diversity pattern reflected with both calculation methods (NMDS and Jaccard-based UPGMA analyses) the observed different community compositions (Fig. 12 and Suppl. Fig. 4A) and explained the separation of protistan communities according to their

structures into a marine-brackish and a brackish group with salinity as the strongest selection factor (Tab. 5) confirming the salinity barrier between 10 – 12 ‰. Likewise, other environmental factors such as temperature and oxygen were identified as significant variables to explain the observed species turnover along the sampling transect (Fig. 12; Tabs. 4 and 5). Furthermore, the distinct nutrient regimes in the Baltic Sea might contribute to the structuring of microplankton communities, too (Filker et al., 2019). However, among these environmental factors, salt is known as one of the strongest environmental barriers to cross for microbes, animals and plants alike (Logares et al., 2009; Forster et al., 2012; Filker et al., 2019). Concluding, the transition zone from 10 ‰ to 12 ‰ represents a strong salinity barrier and structures protistan communities.

After detection of this significant salinity barrier resulting in a protistan plankton community shift and the validation of an increasing alpha-diversity with decreasing salinity in the Baltic Sea, the protistan community structures were compared to hypersaline samples to set in relation whether protistan plankton from brackish waters differ from the structure of protistan plankton found in hypersaline environments. As a consequence, significant shifts in protistan community structures occurred due to different community compositions (Filker et al., 2017) and five different protistan community clusters, namely brackish (8 – 10 ‰), brackish – marine (12 – 22 ‰), marine low-hypersaline (39 – 90 ‰), mid-hypersaline (120 – 240 ‰) and extreme-hypersaline (> 270 ‰), were defined assuming salinity-dependent transitions boundaries in between (Fig. 14). The former proclaimed additional transition boundary between 320 and 330 ‰ by Filker et al. (2017) dividing the extreme-hypersaline cluster in two sub-cluster, cannot be comprehended because of very low bootstrap values (< 0.5) and not clearly differing taxonomic richness (Fig. 14B). Furthermore, the five identified protistan community cluster do not concur with salinity classes reported for bacteria (Javor, 1989: 35 – 100 ‰, 100 – 150 ‰, 150 – 320 ‰; Oren, 1999b: 32 – 100 ‰, 100 – 200 ‰, > 200 ‰; Filker et al., 2017). Nevertheless, salinity appeared to be a stronger selection factor for the structuring of protistan communities than geography (Filker et al., 2017), which was also confirmed in surface water layers of the Baltic Sea by a strong Pearson correlation of beta-diversity with salinity and a negligible relationship of beta-diversity with geographic distance (Fig. 13 and Suppl. Fig. 5). Similar results about the pronounced structuring effect of salinity were even reported

from bacterial communities (Lozupone and Knight, 2007; Herlemann et al., 2011; Dupont et al., 2014) as well as previous studies of protists from other marine-hypersaline habitats (e. g. Forster et al., 2012).

The occurring contradiction between the alpha-diversity measurements of the Baltic Sea based on Simpson index as well as effective number of species (Fig. 10) and the taxonomic richness from the comparison to hypersaline samples (Fig. 14) can be solved by consideration of the different calculation methods. The taxonomic richness measures only how many different OTUs (taxa) occurred neglecting their abundances but the Simpson index and effective number of species weight also the evenness of the OTUs (taxa) in the community. Therefore, the Simpson index and the effective number of species are more reliable dimensions for alpha-diversity because OTUs were not equally distributed within a community (Fig. 11) and the former proclaimed protistan species maximum in the horohalinicum of the Balic Sea by Telesh et al. (2011) can be confirmed (Fig. 10). Nevertheless, the OTU richness did not increase with decreasing salinity (Fig. 10) and hence, a lower taxonomic richness for protistan communities of brackish waters compared to marine and hypersaline environments occurs (Fig. 14B).

The physiological adaptations of organisms to different salinities, which governed the distribution of different taxon groups along a salinity gradient and thus, might be responsible for the formation of transition boundaries and protistan community shifts, remain largely unknown (Filker et al., 2019). First molecular studies show versatile cellular responses of different protists to salinity shifts (Harding et al., 2016; Skarlato et al., 2018; Weinisch et al., 2018b; Weinisch et al., 2018a; Filker et al., 2019). One central strategy is the exclusion of mainly sodium ions from the cytoplasm while acquiring or synthesizing concentrations of organic compatible solutes. Such solutes counterbalance the extracellular osmotic pressure and therefore, maintain the essential cell turgor by minimizing water loss (Yancey et al., 1982; Filker et al., 2019). Due to the investigated halophilic protists *Cyclidium glaucoma*, *Euplotes* sp., *Fabrea salina* and *Pseudocohnilembus persalinus* (Weinisch et al., 2018a), the physiological explanation for the development of a salinity barrier between 10 and 12 ‰ could be provided by the usage of the 'low-salt – organic-solutes-in' strategy because these species accumulate glycine betaine and ectoine as their main osmoprotectants and have their lower salinity tolerance around 12 ‰. Consequently, they require as 'low-salt – organic-solutes-in' strategists a certain minimum of salt and would not survive in the brackish community

cluster. In this study, however, especially dictyochophyceae and dinophyceae were remarkably influenced by salinity shifts, leading to a notable species turnover and selection at salinities below 10 ‰. Thus, these two taxon groups are interesting candidates for more detailed investigations to better understand the cellular mechanisms governing the transition of protists especially from freshwater to marine habitats (Filker et al., 2019). However, the reason for the observed significant community shifts in hypersaline environments reflecting the existence of certain salinity-dependent transition boundaries is unaffected. For the transition boundary between mid-hypersaline and extreme-hypersaline communities, Filker et al. (2017) supposed physiological reasons as the usage of different haloadaptation strategies are the most likely explanation. Oren (2008) explained for 'high-salt-in' strategists that they cannot survive in lower salinities because their highly acidic proteomes will denature at lower salinities (obligate halophiles; Filker et al., 2017). Additionally, organisms relying exclusively on the 'low-salt – organic-solutes-in' strategy to compensate osmotic pressure might reach their limit at this salinity boundary and organisms, which can use both strategies, might be capable of crossing this boundary and could be shared between the two different communities (Filker et al., 2017). Unfortunately, scientific knowledge about metabolic limitations that prevent many protists from crossing certain salinity barriers is insufficient (Harding et al., 2016; Filker et al., 2017). Therefore, it can only be hypothetized that the organisms' ability to use certain haloadaptation strategies is responsible for the different salinity tolerances and different distributions along salinity gradients.

In the era of climate change, it is likely that salinities of aquatic ecosystems will change due to increased precipitation or drought periods (IPCC, 2013). Therefore, knowledge about community shifts and their underlying environmental barriers is essential to predict the possible effects of environmental changes. In the Baltic Sea, for example, a predicted increased precipitation (IPCC, 2013) resulting in lower salinities (Meier et al., 2006) could shift phytoplankton blooms and this would also affect numerous other trophic levels (Sommer et al., 1986; Barber, 2007; Sommer et al., 2012; Filker et al., 2019). One prominent example that even harmful algae blooms could be affected, is the potentially toxic alien dinophyceaen *Prorocentrum minimum* (Kimor et al., 1985; Pertola et al., 2005), which could outcompete its native congener *P. balticum* (Telesh et al., 2016; Filker et al., 2019). *P. minimum* is adapted to salinities above 10 ‰

and has the capability for short-term physiological adaptation to the critical salinity zone (Skarlato et al., 2018). Therefore, it was not suprising to detect the highest relative abundances of this dinophyceaen in the western Baltic Sea within the brackish – marine community cluster and only very low relative abundances in the eastern Baltic Proper (eastern Arkona, Bornhom and Gotland Basins; Filker et al., 2019). But whether the expected salinity changes will favour one of these two species are still open and a field for future research.

In summary, salinity is one of the major driving forces to structure protistan plankton communities even in hypersaline as well as low-salt habitats resulting in specific salinity-dependent transition boundaries. Besides the observed transition boundaries in hypersaline environments (Filker et al., 2017), two additional boundaries are detected in the transition zone from freshwater to marine habitats. The protistan diversity increases with decreasing salinity in brackish habitats as well as in hypersaline communities and the highest protistan diversity can be found in mid-hypersaline habitats. However, the physiological reasons for the observed shifts in community composition along this pronounced salinity gradient are still unknown. But it could be speculated that the usage of different haloadaptation strategies by different organisms may possibly explain these shifts.

2. Analyses of gene expression patterns of microeukaryotes along a salinity gradient

The results from subject 'I' revealed significant shifts in protistan community composition along low-brackish to extreme-hypersaline conditions occurring at specific salinities. With the following environmental, comparative meta-transcriptome analyses of natural protistan communities thriving between 40 and 380 ‰ salinity, the hypothesis was tested whether observed protistan distribution patterns are due to the requirement of different adaptive mechanisms to cope with high salt concentrations. Therefore, the observed shifts in protistan community composition were tried to be linked to different haloadaptation strategies such as the 'high-salt-in' or the 'low-salt – organic-solutes-in' strategy. Alternatively, it was investigated whether it is likely that protists use a combination of both as 'hybrid' strategy, which was found in some bacteria and assumed for a few fungi (Prista et al., 2005; Ramos, 2005; Saum and Müller, 2008; Hänelt and Müller, 2013). The investigation of potential physiological functions and their expression differences between whole protistan communities in response to salinity without cultivation-biases was based on 1,165,559 orthologous groups (OGs).

Only 327,164 (28 %) of the 1,165,559 OGs could be functional annotated. Other studies focusing on metatranscriptomes of marine microbial communities provided similar amounts of annotation (Gilbert et al., 2008; Gifford et al., 2011) and because of the decreasing proportion of functional annotated OGs with increasing salinity (Fig. 15 and Fig. 16), it has to be concluded that suitable databases of hypersaline organisms are still missing and each unannotated sequence of this study could be a starting point for future investigations.

Even though a minority of OGs could be annotated, dinophyceae, ciliophora and chlorophyta are the keyplayers contributing most to the metabolic potential (Fig.16). Compared to literature data, these findings are not surprising because these taxa were known as the main dominant groups at the sampled salinities (Filker et al., 2015; Vargas et al., 2015). For example, dinophyceae are very common in marine environments and they can cause 'red tides' in the oceans (Taylor et al., 2008; Vargas et al., 2015). For ciliophora instead, Filker et al. (2015) reported that ciliophora is the most diverse taxon group at intermediate salinities and chlorophyta have the greatest diversity even at higher salinities. Although the results of this study were based on the taxonomic assignment of proteins and not taxonomic marker genes, chlorophyta dominate the

microeukaryotic communities between 200 and 380 ‰ salinity too, corresponding to microscopy studies of the genus *Dunaliella* in high salt environments since a century (Oren, 2005). For hypersaline environments in general, Filker et al. (2015) demonstrated a decreasing protistan diversity with increasing salinity. Interestingly, the same trend could be observed in this study if the phylum numbers of the different salinities were compared to each other (Fig. 16). Because this study is based on transcribed proteins instead of marker genes, a markedly change of dominant taxa together with a decreasing diversity with increasing salinity suggest that even the used proteins have to differ significantly between different salinities. Due to the fact that the abundance of a protein changes in dependence of salinity, it might be possible that different haloadaptation strategies are used from different taxa at certain salinities.

To test this assumption and to investigate functional (dis)similarities between the samples, the protistan communities from different salinity samples were clustered on the basis of their occurring OGs and their correspondend OG-abundance. On basis of the OGs as the best possible resolution, three main clusters were observed supported by a robust tree topology (Fig. 17): 'marine' (40 ‰ salinity), 'mid-hypersaline' (110 – 150 ‰ salinity) and 'extreme-hypersaline' (200 – 380 ‰ salinity). This metabolic-based cluster pattern reflects the detected transition boundary between 90 and 120 ‰ salinity of subject 'I' separating the marine from the mid-hypersaline cluster. But on basis of the OGs and their abundances, the protistan community shift between 240 and 270 ‰ is not justified because the extreme-hypersaline cluster included 200, 240 and 300 ‰ salinity samples (see Fig. 14 and Fig. 17). Nonetheless, the cluster pattern of protistan communities (Fig. 14) and the OG-based clustering (Fig. 17) are very similar because the OG-based marine cluster reflects the marine-low-hypersaline protistan communities, the OG-based mid-hypersaline cluster is equal to mid-hypersaline protistan communities and only the exact salinity for the separation of the mid-hypersaline from the extreme-hypersaline cluster differs minimal (200 ‰ instead of 270 ‰ salinity).

To identify relevant OGs or metabolic properties being responsible for the observed formation of the three main clusters (Fig. 17) and to ascertain relevant differences in the usage of different haloadaptation strategies by certain protistan communities to explain the observed shifts of subject 'I', all OGs were mapped to the KEGG database for doing a KEGG-based quantitative metabolic fingerprinting.

Unfortunately, neither a KEGG module nor a specific KEGG sub-module show a clear trend for expression value shifts among the three main clusters marine, mid-hypersaline, extreme-hypersaline and fluctuations happen quite frequently. This finding demonstrated that potential osmotically relevant strategies to tolerate different salinities might be masked by the too low and rather general classification level of KEGG. Therefore, focussing on a higher resolution was necessary. To identify osmotic relevant processes and intracellular reactions of the different protistan communities, the annotated functions of the 44,045 differentially expressed OGs were manually grouped to the possibly relevant osmotic categories 'sensing', 'compatible solute transport', 'compatible solute metabolism', 'ion transport' and 'energy'. Using these categories, an attempt was made to explain the observed partitioning of functional diversity.

The revealed three distinct groups of the category 'sensing' (Fig. 19) match perfectly the previously defined marine cluster (40 ‰), mid-hypersaline cluster (110 – 150 ‰) and extreme-hypersaline cluster (200 – 380 ‰). The 70 different functional groups based on 969 manually ordered OGs belonging to the category 'sensing' highlighted significant abundance differences between the three main clusters (Fig. 20). In general, ignoring possible activation or inhibition processes of e. g. enzymes or transporters and assuming that no post-transcriptional or post-translational modifications occurred, the detected amount of transcripts in a community (e. g. TPM values of the OGs) is used as dimension for the amount of translated proteins with the correspondend annotated function in that community. Because mRNAs are not functional entities of a cell, gene functionality can therefore only be inferred indirectly from transcriptomics (Oliver, 2000; Weinisch et al., 2018a). In the following, this means that a higher TPM value for a certain OG at the respective salinity implies a higher capacity of the annotated function of this OG in the community compared to other communities. Within the category 'sensing', the 40 ‰ salinity sample contains the fewest portion of different functions and the mid-hypersaline cluster (110 – 150 ‰ salinity) contains the most different functions decreasing slightly for samples from higher salinities (Fig. 20). Sensing mechanisms and sensing pathways are necessary for organisms to detect changes of their environmental conditions like a salinity shift or to communicate between single cells (Lai, 2004). The general pattern that most pathways were found in the mid-hypersaline cluster indicates that sensing is most important at intermediate salinities. Maybe in the mid-hypersaline cluster occurred both types of

organisms, on the one hand organisms, which live at their maximum tolerance border for salinity and others, which tolerate even higher salinities. This assumption is supported by proteins of osmotic avoidance being detected only in the mid-hypersaline cluster (Fig. 20) illustrating that organisms try to avoid additional osmotic stress. Considering the greatest taxonomic richness occuring at mid-hypersaline salinities (Fig. 14) and organisms need sensitive and secure detection mechanisms for salinity changes, it appears highly likely that the most different sensing functions were observed in the mid-hypersaline cluster as a higher species diversity could direct to a higher diversity of sensor functions.

Exemplary, the two component response showed the highest capacity for the mid-hypersaline cluster with decreasing capacity at higher salinities and missing completely for the marine cluster (40 ‰). Similar results were observed for mitogen-acticated protein kinases (MAPK) and hybrid signal transduction histidine kinases (Fig. 20). Normally, the two component signal transduction system consists of a sensor histidine kinase in cooperation with a response regulator (Capra and Laub, 2012) illustrating one of the simplest and effective signal transduction pathways in all three domains of life (Wuichet et al., 2010). For eukaryotes under osmotioc stress, the two component signal transduction system, which was mainly studied in yeasts (Maeda et al., 1994), includes a hybrid signal transduction histidine kinase, followed by a histidine phosphotransfer protein and a response regulator containing a receiver domain before a MAPK cascade is activated (Loomis et al., 1998; Suescún-Bolívar and Thomé, 2015). The higher capacity of the components for the two component signal transduction system in hypersaline samples compared to the marine cluster (Fig. 20) illustrates the importance for organisms to detect the extracellular salinity. Their need to response could be observed by means of the category stress response representing the lowest capacity at 40 ‰ and increasing constantly with increasing salinity until the highest capacity was detected for communities in the extreme-hypersaline cluster (Fig. 20).

The increasing capacity for the thioredoxin-based redox signalling with inceasing salinity (Fig. 20) potentially evidenced that those organisms have to deal not only with osmotic but also with oxidative stress at high salinities. This happens because with increasing salinity the solar saltern ponds become shallower and shallower resulting in a higher UV radiation dose for each organism and oxidative stress conditions could increase (Toone et al., 2001). The same finding was observed for the universal category

stress response showing the highest capacity in the extreme-hypersaline cluster (Fig. 20) indicating the most stressful conditions for organisms occur at higher salinities.

In contrast, the slightly decreasing capacity of calmodulin with increasing salinity showing the highest capacities for marine communities was surprising. Calmodulin is a small protein with highly conserved amino acid sequence throughout evolution tolerating heat and acids (Klee et al., 1980). Calcium signalling via calmodulin is a widely distributed process to regulate the stress response of a large number of different enzymes (Klee et al., 1980; Kraus and Heitman, 2003). Therefore, the calmodulin based Ca^{2+} signalling should play an important role in all protistan communities. This could be concluded from the relatively high capacity in all samples (Fig. 20). But calmodulin might be also relatively unspecific in case of osmotic stress and thus the slight decrease with increasing salinity could be explained. Additionally, higher capacities of possible Ca^{2+}/calmodulin targets (Fig. 20) were found in hypersaline samples assuming that the Ca^{2+}/calmodulin targets might be responsible for a more specific cellular reaction to osmotic stress. In summary, several different sensing mechanisms and pathways were responsible with their unequal capacities in different protistan communities for the observed cluster pattern (Fig. 19) concluding that sensing, especially osmosensing, is a partly specialized process and essential for organisms living at high salinities.

To link potential changes in the usage of different haloadaptation strategies with the observed protistan community shifts (subject 'I'), the detection of compatible solutes and intracellular ions is essential. Hence, the transport and the synthesis of possible compatible solutes will be focused first followed by the investigation of potential ion transport capacities within the different protistan communities. The category 'compatible solute transport' provided with its annotated functions from 113 OGs, resulting in 26 presumed different transport mechanisms for at least 21 potential compatible solutes (Roberts, 2005; Empadinhas and da Costa, 2008), a very similar cluster pattern as it was shown for all OGs (Fig. 17) forming a marine, a mid-hypersaline and a extreme-hypersaline group (Fig. 21). Most putative transporters were detected in the mid-hypersaline cluster (110 – 150 ‰ salinity) indicating a higher diversity of potential compatible solute transport mechanism at these salinities. Due to the greatest taxonomic richness in the mid-hypersaline cluster (Fig. 14), the higher quantity of presumed different compatible solute transporter could result from the higher amount of different organism using differential transport mechanisms. Comparing the

abundances of the different putative transporters according to salinity and thus, receiving variable transport capacities within the certain protistan communities, it was obvious that the potential compatible solute transport capacities were at its maximum in mid-hypersaline samples, especially for the compatible solutes choline and trehalose (Fig. 22). In contrast, the putative capacity for sugar transport increased fundamentally from 40 ‰ to the mid-hypersaline cluster and reached its maximum in the extreme-hypersaline group (Fig. 22). Supposed transport capacities for glycine betaine or ectoine were only detected in marine and mid-hypersaline samples but missing completely in the extreme-hypersaline cluster (Fig. 22). For these reasons, it could be assumed that the usage of compatible solutes to counterbalance the osmotic stress of the surrounding environment is very common in the mid-hypersaline cluster, almost reduced in the marine sample and an occurring but not dominant process in extreme-hypersaline communities explaining the observed cluster pattern (cf. Fig. 21). Concerning compatible solutes, e. g. ectoine, could also be used as an energy resource containing carbon and nitrogen (Czech et al., 2018), and sugars are predominantly energy-related molecules, it is very likely that the higher presumed capacity for unspecific sugar transport in extreme-hypersaline samples was motivated by a higher energy demand of the organisms living in higher salinities (Oren, 1999a). Consequently, it has to be concluded that higher transport capacities imply not mandatorily that the transported possible compatible solutes were really accumulated and used for osmoprotection.

To investigate whether a potential compatible solute is used for osmoprotection or as energy resource instead, the expression values of possible enzymes for the *de novo* synthesis or the catabolism of 17 known compatible solutes were compared (Fig. 24). The UPGMA cluster pattern indicated a significant difference between marine and hypersaline communities and the establishing of a mid-hypersaline cluster could only be assumed on the basis of the tree topology but missing convincing bootstrap support (Fig. 23). This cluster pattern is reflected by presumed lower synthesis as well catabolic capacities of all possible compatible solutes in the marine sample and very high capacities in hypersaline samples (Fig. 24) concluding that both synthesis and catabolism of potential compatible solutes might be minor dominant in the marine cluster. To identify molecules, that could be accumulated as compatible solutes in the different protistan communities and those, which might function as energy resource or be need for other metabolic processes, the presumed capacities for the *de novo* synthesis

and the catabolism were compared based on the expression of the respective known enzymes. Because of a relatively high capacity for both synthesis and catabolism of glutamate in all hypersaline samples (Fig. 24), it is verly unlikely that glutamate is accumulated and used as compatible solute in the whole protistan community and thus, glutamate serves probably as essential amino acid in other cellular processes. For the reason that meta-transcriptomic data includes many different species, it is however still possible that some organisms use glutamate as compatible solute and other organisms within the same community take glutamate as energy resource, explaining the putative high capacity of degrading enzymes with simultaneous high synthesis capacity. In contrast, the presumed synthesis capacity for *myo*-inositol, trehalose, sucrose and proline potentially overmatched the degradation capacity especially in hypersaline samples (Fig. 24). Therefore, it can be assumed on community level that these molecules are primarily used as compatible solutes in hypersaline protistan communities. Additionally, for hydroxyectoine and glycine betaine no degrading enzymes were detected and because of their assumed synthesis capacity it can be reasoned that hydroxyectoine functions as compatible solute mainly in the mid-hypersaline cluster and glycine betaine additional in the extreme-hypersaline communities. Interestingly, for both precursors of hydroxyectoine and glycine betaine, namely ectoine and choline, no clear pattern, neither for synthesis nor for catabolism, was observed due to the salinity-dependent community clustering and thus it might be concluded that both are not primarily used as compatible solute in a specific community. Likewise, Czech et al. (2018) reported that especially ectoine is used by many different organisms as a high quality carbon and nitrogen resource benefitting from other species within the same community, which produce ectoine abundantly for osmoprotection. Nevertheless, other authors proclaimed that ectoine is one of the preferred compatible solute in low and mid-hypersaline environments while hydroxyectoine enables species to tolerate even higher salinities (Harding et al., 2016; Czech et al., 2018; Weinisch et al., 2018a). Referring to this hypothesis, a higher synthesis or import capacity for ectoine in the mid-hypersaline cluster and for hydroxyectoine in extreme-hypersaline communities was assumable but could not be verified by the findings of this study.

Considering the results for glycerol and aspartate that their putative catabolic capacity exceeded their putative synthesis capacity especially in hypersaline samples, it is unlikely that both might be accumulated as compatible solutes. For asparagine

instead, the presumed synthesis capacity was overbalanced in the marine and extreme-hypersaline cluster while supposed catabolism and synthesis capacities were balanced in mid-hypersaline communities (Fig. 24). In conclusion, asparagine might only function as compatible solute at high salinities, if at all. In summary, the findings of the different putative transport as well as catabolic or synthesis capacities for selected potential compatible solutes provide the conclusion that compatible solutes are accumulated for osmoprotection in hypersaline environments, especially predominant in the mid-hypersaline cluster. But a trusted exclusion of the 'low-salt – organic-solutes-in' strategy from one of the three main clusters can not be guaranteed. It is more likely that compatible solutes accumulating species occur in all investigated protistan communities.

To test whether the same picture can be drawn for the 'high-salt-in' strategy and to evaluate whether a mixture or 'hybrid' strategy is common in protistan communities, the potential ion transport capacities within the different communities were analysed. Typical ion transport mechanisms involved in ion homeostasis and osmoadaptation are known from yeasts and selected fungi (Prista et al., 2005; Ariño et al., 2010; Ramos et al., 2011; Zajc et al., 2013). For example, the reviews of Ariño et al. (2010) and Ramos et al. (2011) reported that K^+ influx is mostly realised in yeasts by transporters of the Trk (Transport of K^+) family and K^+ efflux is guaranteed by Tok channels together with also Na^+ expulling transporters of the Nha family. Another exclusively in fungi, bryophyte and protists detected either Na^+ or K^+ expulling transport mechanism are the ATPases of the Ena (Efflux of Na) family, reviewed by Rodríguez-Navarro and Benito (2010). Additionally, the yeast *Debaryomyces hansenii* possesses a protein homolog to the Nhx transporter of *Saccharomyces cerevisiae*, which could be responsible for the increased halotolerance of *D. hansenii* via Na^+ sequestration into the vacuole (Prista et al., 2005). On the basis of the 336 manually grouped OGs representing 46 different putative ion transport mechanisms, a reliable division could not be conducted between transporters or transport mechanisms, which are relevant for universal cellular ion homeostasis, and others, which are specialized for salt adaptation. While the universal ion homeostasis is based on a constant Na^+/K^+ ratio and an increasing salinity provides an imbalance due to an increased Na^+ level, all organisms require transport mechanisms for the enhancement of K^+ and the intracellular reduction of Na^+. For that reason even 'low-salt – organic-solutes-in' strategists possess ion transporters like Trk, Nha, Ena or Nhx

(Prista et al., 1997; Kogej et al., 2005; Zajc et al., 2013; Plemenitaš et al., 2014). Therefore, the higher putative ion transport capacities in hypersaline samples compared to the marine community are not suprising and indicate a higher relevance of ion transporting mechanisms in protistan communities of hypersaline environments for both, ion accumulation and ion homeostasis (Fig. 26). More precisely, the highest diversity of different putative ion transport mechanisms and the greatest presumed transport capacities for inorganic ions at salinities of 110 and 150 ‰ (Fig. 26) indicate that both processes, ion accumulation and ion homeostasis, could happen in the mid-hypersaline cluster. Due to the fact that 'low-salt – organic-solutes-in' strategists are sensitive to an increased amount of intracellular Na^+, the expulsion capacity for mainly Na^+ was important for the evaluation whether the 'low-salt – organic-solutes-in' strategy is used in certain protistan communities. According to the mentioned Nha and Ena transporter families, Na^+ expulsion might be realised by the detected cation-transporting ATPases, Na^+/K^+-transporting ATPases or Na^+/H^+ antiporter (Ariño et al., 2010; Ramos et al., 2011). These transport mechanisms were dominantly found in samples belonging to the mid-hypersaline cluster and because of the slight decrease of their assumed capacity in extreme-hypersaline samples, the 'low-salt – organic-solutes-in' strategy is predominantly used in the mid-hypersaline cluster and not so common in extreme high salinities (Fig. 26). Other examples for possible transport mechanisms belonging to adaptation strategy-independent ion homeostasis were found with anion transporter, voltage-dependent anion channels, vacuolar Ca^{2+} transporter, Na^+/sulfate co-transporter, Cobalt/Mg^{2+} transport and Ca^{2+}-transporing ATPases because their assumed capacities varied independently from the salinity of the sample (Fig. 26). In contrast, the possible realisation of a 'high-salt-in' strategy could be concluded from the the import capacities of Mg^{2+}, K^+ and Cl^- because these ions are almost exclusively used for intracellular ion accumulation instead of harmful Na^+ (Médicis et al., 1986; Oren, 2013) by 'high-salt-in' strategists (Oren, 2008; Gunde-Cimerman et al., 2018). Additionally, the presumed unspecified ion transporter, which showed an increasing capacity with increasing salinity until attaining its maximum at 240 ‰ salinity (Fig. 26), could function in osmotic stress independent ion homeostasis as well as for intracellular ion accumulation. However, the 'high-salt-in' strategy might be realised in this dataset possibly by Mg^{2+} transport, H^+/Cl^- exchanger or Ca^{2+}-activated K^+ channels. For example, it is conceivable for the Ca^{2+}-activated K^+ channels that Ca^{2+} functions as signal transmitter (Clapham, 2007) to perceive osmotic stress and the activation of K^+ channels

could represent the cellular reaction to counterbalance this stress via an increased K^+ influx (cf. Li et al., 2006). Due to the detected different capacities of Mg^{2+} transport, H^+/Cl^- exchanger and Ca^{2+}-activated K^+ channels (Fig. 26), it could be assumed that the usage of the 'high-salt-in' strategy might be realized mainly via K^+ accumulation in the mid-hypersaline cluster and preferred Mg^{2+} accumulation in extreme-hypersaline samples. But a constantly high capacity of K^+ efflux antiporter in all hypersaline communities might indicate also the usage of alternative strategies, like 'low-salt – organic-solutes-in', to deal with high salt concentrations. As a consequence, the 'high-salt-in' strategy is more likely used in extreme-hypersaline communities. To verify whether a detected transport mechanism is involved in a specific haloadaptation strategy or in universal ion homeostasis, direct ion flow measurements and more detailed analyses of the single cell will be challenges for further studies in the future.

Regarding the varying energy demands for the realisation of different haloadaptation strategies (Oren, 1999a; Oren, 2001) and if protistan communities living at different salinities differed according to the usage of their main haloadaptation strategy, it should be possible to identify this strategy shift with the help of the investigation of energy-related processes because the energy production capacity should vary in dependence of the salinity of the sample. A significant and well supported shift was only observed between the marine cluster and the hypersaline communities (Fig. 27). The 3691 OGs representing 41 different mechanisms or energy-related processes provided a definite result. With the exception of sulfur metabolism, all other investigated processes had putatively an increased capacity in hypersaline samples and a lower capacity at 40 ‰ salinity (Fig. 28). Only the presumed capacity of lipid metabolism increased constantly with increasing salinity reaching its maximum at 380 ‰ salinity (Fig. 28). Concluding from these findings, organisms living in hypersaline environments should have a significant higher energy demand than marine species. Therefore, a clear shift in the usage of haloadaptation strategies could only be reasoned between the marine cluster and hypersaline protistan communities. A more detailed resolution of energy-related processes to find potential variations within the hypersaline group was not possible due to the general necessity of energy production in each living organism (Oren, 2006), symbolized by the basic metabolic categories 'respiration', 'photosynthesis', light harvesting', 'glycolysis', 'carbohydrate metabolism' and 'ATP synthase' (Fig. 28).

Besides the investigation of the assumed energy demand respectively the putative energy production capacity of a protistan community, the acidity of the cytoplasmic proteome of a species reveals also whether the 'high-salt-in' or an alternative strategy is used. Because the 'high-salt-in' strategy is characterized by the accumulation of inorganic ions inside the cell, mainly K^+, the enzymatic machinery has to be adapted to such high intracellular ion concentrations by an excess of acidic amino acids on their protein surface (Paul et al., 2008; Rhodes et al., 2010). Therefore, if the enzyme machinery tolerates high intracellular ion concentrations, proteins contain more acidic amino acids and the whole proteome is enriched with acidic proteins. With the calculation of the isoelectric points of cytoplasmic proteins and the comparison between different salinities, the average acidity of all cytoplasmic proteins occurring in the specific protistan communities was taken as indicator whether the majority of the community has a well-adapted cytoplasm for tolerating higher intracellular ion concentrations or not. Although the statistical comparison provided significant shifts in the averaged acidity for the cytoplasmic proteome of the different protistan communities (letters in Fig. 29), no distinct difference in matters of salt tolerance for the certain communities could be observed considering the overlapping of the 95 %-confidence intervals between all salinities (borders of the boxplots in Fig. 29). Even an additional separation of the cytoplasmic proteins according to their taxonomic affiliation achieved no significant and biological relevant shifts in case of proteome acidity and intracellular salt tolerance (Suppl. Fig. 7).

In summary, changes of metabolic properties and putative capacities within protistan communities along a salinity gradient identified three main clusters. Specific haloadaptation strategies should be minor dominant in marine communities comparing the capacities of all investigated osmotic relevant processes between the marine cluster and hypersaline samples. The presumed different transport capacities of compatible solutes and ions split hypersaline communities into two cluster, mid-hypersaline and extreme-hypersaline. According to these capacity differences and assuming that synthesis and catabolism of a compatible solute could occur within the same community but in different organisms, it is supposable that the 'low-salt – organic-solutes-in' strategy is dominantly used for osmoprotection in the mid-hypersaline cluster and sporadically in the extreme-hypersaline cluster. Furthermore, the usage of the 'high-salt-in' strategy cannot be rejected and might even increase with increasing salinity because

specific putative ion transporters and channels for influx and accumulation of Mg^{2+}, K^+, Cl^- were detected, besides transport mechanisms for universal ion homeostasis. The question remains unanswered whether a 'hybrid' strategy or other forms of both main strategies are common in protists although it might be possible for some organisms. Concluding the usage of the 'high-salt-in' strategy is more probable in communities at higher salinities and the 'low-salt – organic-solutes-in' strategy might be preferred in mid-hypersaline communities, different protistan community compositions can be linked to certain haloadaptation strategies but an entire separation of protistan communities regarding the usage of different haloadaptation strategies cannot be verified.

3.   Time-resolved transcriptome analysis of *Schmidingerothrix salinarum*

Considering the former results of subject 'I' and subject 'II', shifts in protistan community composition occurred in dependence of salinity and the 'low-salt – organic-solutes-in' strategy might be preferred by protists tolerating salinities up to 150 ‰, detailed and comprehensive informations about involved enzymes, specific pathways and cellular processes associated with the 'low-salt – organic-solutes-in' strategy were not obtained. Therefore, a time-resolved transcriptome analysis of the model ciliate *Schmidingerothrix salinarum* was conducted to investigate the organism's expression dynamics after osmotic up-shock and during salt acclimation. With this approach possible cellular reactions to a salt up-shock could be identified with a time-dependent resolution highlighting putative time-dependent intracellular processes of *S. salinarum*.

The received global transcriptome of *S. salinarum* contained 168,454 ORFs representing 48,022 different annotated functions. Because of the nescience of the exact genome size and the number of genes due to the absence of a completely sequenced reference genome of *S. salinarum*, the received amount of ORFs seemed to be too high compared to closely related species. Although ciliates have macronuclei (MACs) for vegetative, clonal growth (Herrick, 1994) and germline micronuclei (MICs) for sexual reproduction (Cavalcanti et al., 2004), their transcribed mRNA originates exclusively from nanochromosomes of the MAC (Prescott, 1994; Cavalcanti et al., 2004). Those nanochromosomes contain almost nearly pure coding DNA plus short regulatory sequences and are highly amplified by polytenization of MIC chromosomes during differentiation (Doak et al., 2003). For example, *Oxytricha trifallax* possesses approximately 24,500 different nanochromosomes carrying 26,800 genes (Prescott et al., 2002; Doak et al., 2003) and for the close relative *Stylonychia lemnae*, Aeschlimann et al. (2014) assembled its macronuclear genome receiving more than 16,000 nanochromosomes and 19,851 contigs. Eisen et al. (2006) predicted for *Tetrahymena thermophila* more than 27,000 protein-coding genes and Aury et al. (2006) reported for *Paramecium tetraurelia* 39,642 genes. Because of the different numbers of MICs and MACs within the phylum ciliophora, the number of genes and gene copies vary. For example, Tetrahymena species have one MIC and one MAC, Paramecium species have 2 or more MIC and one MAC and Stylonychia as well as Oxytricha species have two MACs and 2 or 2 to 4 MICs (Prescott, 1994). With regard to the well documented number of isoforms and varying copy numbers in the phylum ciliophora (Herrick, 1994; Prescott,

1994; Jahn and Klobutcher, 2002) and the observations of Weinisch et al. (2018b) that *S. salinarum* possesses one MIC but six MACs resulting in many gene copies, the received amount of 168,454 ORFs could be explained by combining a lot of gene copies with similar functions originated from 19,000 – 21,000 extracted imdividual cells per replicate. Thus, it could be assumed that the represented 48,022 different annotated functions described a more realistic genome size of *S. salinarum*. But the actual genome size and its gene number can only be defined by genome sequencing of *S. salinarum*`s MACs in the future.

The completeness of the used global transcriptome to receive 15,093 differentially expressed ORFs representing 9,239 different gene functions, was checked by searching the KEGG annotations for essential metabolic pathways and structural complexes (Beisser et al., 2017). For all datasets, the annotations of KEGG provided 21 pathways and complexes, which were almost complete and whose central modules were always functioning (Fig. 30). Together with the concruency of all three replicates at each time point shown in the sample correlation heatmap (Fig. 31), a satisfied completeness and reliability for the global transcriptome was reached. Additionally, the TMM-normalized expression counts per ORF illustrated with the sample correlation heatmap a cluster pattern for the different time points (Fig. 31), which could be a hint for the execution of certain intracellular processes at different time points after salt up-shock.

This hint was substantiated by clustering the differentially expressed 15,093 ORFs according to their expression compared to t0. The uniform reference t0 was chosen to receive time- and salinity-dependent results because all regulatory changes at each time point were compared to a consistent reference, whose cells were not osmotically shocked. Therefore, all received changes can be linked directly to the increased salinity as trigger. Thus, it can be concluded from the received time-dependent pattern of down- and up-regulated ORFs (Fig. 32) that the cell requirements varied over time after an osmotic up-shock. Furthermore, the sample correlation showed two distinct sample groups besides control. Consequently different phases of adaptation might be possible, namely samples 2 min, 10 min, 60 min and 120 min (group 1) versus samples 30 min and 720 min (group 2).

To investigate whether characterizing metabolic processes or intracellular reactions could be identified for the certain phases or for particular time points, the

KEGG-annotated 3,202 up- or down-regulated ORFs were ordered according to their KEGG categorization. Additionally, all differentially expressed ORFs were grouped according to their sub-cellular localisation to test whether the localisation of molecules played an important role during osmotic adjustment and to reveal compartments, where the cellular response to the salt-up shock was most active, because ion sequenstration could happen in *S. salinarum* to survive under stressful osmotic conditions indicated by ion imaging experiments of Weinisch et al. (2018b). Comparing the proportion of up- and down-regulated ORFs within the same KEGG category according to the different time points after salt up-shock, differences were observed. Within the category 'genetic information processing' explicitly more ORFs were up-regulated for 'group 1' than for 'group 2' and the highest proportion of down-regulated ORFs was detected after 30 min (Fig. 33). Therefore, it is assumable that processes like transcription, translation, folding, sorting and degradation as well replication and repair varied significantly between both sample groups and were more dominant in 'group 1'. The nucleus as the place for 'genetic information processing' confirmed this pattern because most up-regulated ORFs occurred after 2 min and 10 min followed by the largest amount of down-regulated ORFs after 30 min and 60 min (Fig. 34). Additionally, a significant difference in the carbohydrate metabolism can be concluded between the time points 10 min and 30 min because the largest proportion of ORFs were up-regulated after 30 min just after the largest proportion of down-regulated ORFs at 10 min (Fig. 33). The cytoplasm, instead, as critical place for osmotic adjustment and for most reactions showed no clear time-dependent pattern according to down- or up-regulated ORFs (Fig. 34). Therefore, it can be concluded that too many osmotic relevant processes occurred at each time point, which masked a potential site-specific pattern. Furthermore, expression differences were found for 'energy metabolism' and 'amino acid metabolism' at 120 min showing the largest amount of down-regulated ORFs (Fig. 33). Hence, the two categories seem to be less dominant in *S. salinarum* after 120 min during the process of haloadaptation. According to this, most mitochondrial ORFs were down-regulated after 120 min but up-regulated at 30 min before (Fig. 34), which reflects that the main energy production is located in mitochondria (Gray, 2002; McBride et al., 2006). Although the greatest relative proportion of down-regulated ORFs in the category 'environmental information processing' was detected after 2 min and the largest relative amount of up-regulated ORFs was observed after 720 min (Fig. 33), the processes of 'signal transduction', which falls in this category, had the largest relative amount of up-regulated ORFs at 120 min

and the greatest relative proportion of down-regulated ORFs after 720 min instead. Assuming logically that signal transduction should be one of the earliest reactions to detect changes in salinity, it has to be concluded that the categories of KEGG have a too low resolution to highlight the correct reaction time points for specific intracellular processes. However, the cell membrane as locality with direct contact to the environment showed a time-dependent expression pattern. A higher proportion of ORFs was up-regulated at 2 min, 10 min and 120 min compared to 30 min, 60 min and 720 min (Fig. 34). Possible membrane proteins could be sensors, channels, pumps or transporters, which are discussed in detail subsequently to explain this observed expression pattern.

Because of an unsatisfied resolution of KEGG categories and insufficient consideration of only sub-cellular localities, in-depth investigations targeting ORFs specifically relevant for haloadaptation strategies were necessary. Therefore, the 15,093 differentially expressed ORFs were assigned to the osmotic relevant categories 'sensing', 'compatible solute transport', 'compatible solutes' and 'ion transport' to elucidate the intracellular dynamics of *S. salinarum* after salt up-shock in a time-dependent resolution. Former ^{1}H-NMR measurements and lab experiments of Weinisch and colleagues (2018b) revealed that *S. salinarum* is an 'low-salt – organic-solutes-in' strategist accumulating glycine betaine (GB) and ectoine (Ect) as compatible solutes. The reported significant positive correlation between the external salinity and intracellular concentrations of GB and Ect as well as the absence of an increase of cytoplasmic Na^+ concentration with increasing salinity (Weinisch et al., 2018b) illustrated the necessity to detect environmental changes especially in case of salinity and the resulting osmotic pressure. Without a fast and precise detection of salinity changes, *S. salinarum* would not be able to tolerate high salinities by accumulating the correct amount of compatible solutes or by keeping the cytoplasmic Na^+ concentration on a constant level. For this reason, *S. salinarum* requires specific sensing mechanisms. The review of Suescún-Bolívar and Thomé (2015) described the hyper-osmolar glycerol (HOG) pathway, originally characterized in yeasts (Loomis et al., 1998), as the main molecular pathway directly involved in the response to high osmolarity. Commonly, this and homologous pathways can be divided in a sensing and a response part involving the activation of phosphorylation cascades (Maeda et al., 1994). Generally, a two-component histidine kinase system consists of a transmembrane histidine kinase protein forming the sensing

part followed by the response part composed of a histidine phosphotransfer protein and a response regulator having a receiver domain to activate MAPK cascades (Loomis et al., 1998; Suescún-Bolívar and Thomé, 2015).

Interestingly, several components for sensing and cell signalling were found in the transcriptome of *S. salinarum*. Because of their high functional similarity to the mentioned HOG pathway, it could be assumed that *S. salinarum* uses perhaps comparable osmosening mechanisms and similar two-component signal transduction systems (Fig. 35). Assuming that no post-transcriptional or post-translational modifications occurred, it is hypothesized that a measured higher expression represented an up-regulation and directly reflected a higher amount of the specific protein within *S. salinarum* at the explicit time point. Thus, the detected time-dependent expression activity symbolized possibly the usage of osmosensing mechanism in *S. salinarum*. According to the sensing part of the mentioned HOG pathway, an upregulation of oxygen sensor histidine kinases and histidine kinases of a two-component sensing system occurred first and were completed by the up-regulation of other histidine kinases and hybrid sensors after 10 min (Fig. 35). Equally, cAMP-regulated phosphoproteins and cAMP-activated global transcriptional regulators could function as response part due to their simultaneously up-regulation (Fig. 35) but this is speculative on the basis of a transcriptome analysis. Furthermore, the measured transcriptional expression reflects not the exact reaction time of a specific pathway because Hersen et al. (2008) demonstrated *in vivo* that the HOG pathway in *Saccharomyces cerevisiae* needed for maximal phosphorylation of the mitogen-activated protein kinase 'Hog1' only one minute. Consequently, sensing should be also in *S. salinarum* a fast mechanism to react quickly to an increasing salinity. Therefore, the expression patterns of different sensing components at 10 minutes or more after salt up-shock could possibly reveal on the one hand which components and sensing molecules have to be renewed after a certain time or on the other hand which molecules are required in a higher amount due to the higher external salinity. These renewed components could predominantly be mitogen-activated protein kinases (MAPKs) and related MAPK phosphatases or inositol-tetrakisphosphate 1-kinases. Examples for a salinity-dependent higher molecule demand could be potentially concluded from the up-regulation of osmotically inducible proteins, response regulators and their receiver proteins after 720 min (Fig. 35). Additionally, the up-regulation of thioredoxin and

thioredoxin reductases after 120 min and thioredoxin-disulfide reductases at 720 min could possibly symbolize the intracellular redox-sensing and protection mechanism from reactive oxygen species (ROS) (Masutani and Yodoi, 2001; Go and Jones, 2013). Furthermore, a third possibility has to be considered: the influence of technical biases could presumably mask the observed expression results, especially for the dataset of 30 min after osmotic up-shock, because the expression values of two-component system sensor histidine kinase, thymidine kinase, signal transduction histidine kinase, oxygen sensor histidine kinase and cAMP-activated global transcription regulator after 30 min differed unexpectedly compared to 10 and 60 min (Fig. 35). Thus, the biological trueness of this observation is questionable and should be tested by further investigations using quantitative real-time PCR (qRT-PCR).

After detection of the increased salinity, there is a need to balance the osmotic pressure. Due to the findings of Weinisch et al. (2018b), *S. salinarum* accumulated therefore compatible solutes. To answer the question whether *S. salinarum* realises this compatible solute accumulation via import or *de novo* synthesis or if there is a preference for one mechanism or both, Weinisch et al. (2018b) performed uptake experiments of labelled choline, which demonstrated indeed that choline is metabolized into glycine betaine (GB). But it is still unclear whether this conversion was done by *S. salinarum* itself or by associated food bacteria, which were then eaten by *S. salinarum*. As a consequence, only this time-resolved transcriptome analysis might answer this open question. Focussing first on the less energy demanding alternative, nine different putative compatible solute transporters with a time-dependent expression pattern were detected and ordered according to their predicted sub-cellular localisation. Assuming no post-transcriptional activity variations occur and transporter activation or inhibition play a minor role, a transporter up-regulation is equalized with a higher transport capacity. As kind of an early reaction after 2 and 10 min, the choline transport capacity over the cell membrane into the cytoplasm increased and with a time delay of nearly 120 minutes, choline was transported with a higher capacity into the mitochondrion (Fig. 36). Considering that choline can be both a compatible solute and a precursor for GB, it can be assumed that choline is possibly used in the cytoplasm first as compatible solute symbolizing a stress response and then as precursor for the GB synthesis in the mitochondrion for osmotic acclimatisation. Weinisch and colleagues (2018b) detected only GB and Ect as main compatible solutes in well-adapted *S. salinarum* cells via [1]H-

NMR spectroscopy. Additionally, the higher transport capacity for glycine betaine located in lysosomes or vacuoles after 60 and 120 min (Fig. 36) could be interpreted as kind of putative storage accumulation if GB was synthesized from choline before because an up-regulated GB transport mechanism located in the cell membrane was not found. Another possibility could be that the predicted location of the different transporters via machine learning was maybe not correct and the transporters might be integrated elsewhere. But assuming a correct locality prediction, *myo*-inositol and sucrose were probably imported with a higher capacity into the cytoplasm after 30 min (Fig. 36) indicating that perhaps both molecules could be used temporary as compatible solutes if they are present in the surrounding environment. Additionally, the obvious down-regulation of Na[+]/*myo*-inositol co-transporter after 720 min might demonstrate why *myo*-inositol was not measured via ^{1}H-NMR spectroscopy by Weinisch et al. (2018b). Considering each potential compatible solute can either be metabolized in other processes or can function as energy resource, the slightly increased alanine symport capacity as well as the import of UDP-N-acetylglucosamine into lysosomes or vacuoles could be explained by these alternative functions. In contrast, the significantly increased transport capacity for proline into the cytoplasm as well as into lysosomes or vacuoles after 720 min permits the conclusion that proline is presumably used as compatible solute although it was not detected via ^{1}H-NMR before (Weinisch et al., 2018b). Interestingly, besides the discovered putative transport mechanisms for the probable compatible solutes choline, GB, *myo*-inositol, sucrose and proline showing a time-dependent expression pattern, an annotated ectoine transporter and ectoine/proline transporter as well was found in *S. salinarum* showing no capacity variation after salt up-shock (see 'Annotationtable_168454_ORFs' in digital appendix). Possibly, the amount of ectoine in the surrounding environment of *S. salinarum* was not large enough that an increased import capacity might be profitable.

Considering four potential compatible solutes were transported between different cell compartments and partly not primarily imported into the cell, they had to be synthesized before. Investigating enzymatic pathways for synthesis or catabolism of potential compatible solutes, the question whether *S. salinarum* is able to synthesize certain compatible solutes or uses them as energy resource could be answered. Enzymes for the synthesis or catabolism of seven potential compatible solutes with a time-dependent expression pattern were found in the transcriptome of *S. salinarum*.

Comparing the compatible solute specific enzyme expression for synthesis and catabolism with each other, it is not highly likely that *S. salinarum* accumulates glycerol or glutamate as compatible solutes. The annotated catabolic enzymes for glycerol degradation were always up-regulated when the synthesis pathway was simultaneously down-regulated and regarding glutamate, possible enzymes for synthesis and catabolism were balanced (Fig. 37). In contrast, a time-dependent shift in the regulation of synthesis and catabolism for *myo*-inositol was found (Fig. 37). But in consideration of the former observed increased capacity for *myo*-inositol import at 30 min and simultaneous up-regulation of annotated *myo*-inositol-2-dehydrogenase as part of the assumed catabolic pathway, the hypothesis that *myo*-inositol is used temporally as compatible solute, seems to be unlikely although the synthesis enzyme inositol-3-phosphate synthase was up-regulated after 120 min (Fig. 37). Respecting the review of Gillaspy (2011), which demonstrates the role of *myo*-inositol-containing molecules in signal transduction, and considering that also a potential inositol 1,4,5-triphosphate receptor was found in the transcriptome of *S. salinarum*, it has to be concluded that *myo*-inositol is more likely a intracellular messenger than a compatible solute within *S. salinarum*.

Knowing that *S. salinarum* uses GB as compatible solute (Weinisch et al., 2018b) and a possible higher choline import capacity was detected after 2 and 10 min (Fig. 36), the up-regulation of potential enzymes for choline synthesis at 10 and 30 min (Fig. 37) could be interpreted as follows: The cell has a higher demand for choline early after salt up-shock and tries to cover this demand first via increased import. If this import was not sufficient, additional choline synthesis is up-regulated. Due to the parallel up-regulation of choline dehydrogenase (CDH) after 30 min as first step of GB synthesis pathway found in *S. salinarum* (Suppl. Fig. 6), choline seems to appear predominantly as precursor and not as compatible solute itself. The second enzyme for GB synthesis, betaine-aldehyde dehydrogenase (BADH), was up-regulated after 60 min (Fig. 37) indicating the potential peak of GB synthesis after an osmotic up-shock. Another hint for this conclusion is provided by the increased GB transport capacity into the lysosomes or vacuoles (Fig. 36) because a higher transport capacity is only needed when more substrate has to be transported. In conclusion, *S. salinarum* is thus able to synthesize the compatible solute GB via a two step reaction (Suppl. Fig. 6) and imports the precursor choline first, which was supported by former growth and feeding experiments of

Weinisch et al. (2018b). Besides other possibilities for GB synthesis (Rathinasabapathi et al., 1997; Nyyssölä et al., 2000; Fan et al., 2004), similar two-step reactions using CDH and BADH are reported from *Escherichia coli* (Lamark et al., 1991), *Bacillus subtilis* (Boch et al., 1996), *Halomonas elongata* (Cánovas et al., 2000) and *Halobacillus dabanensis* (Gu et al., 2008).

Interestingly, Weinisch et al. (2018b) measured for the first time ectoine (Ect) functioning as compatible solute within a protist and Harding et al. (2016) found all ectoine biosynthesis genes (EctB, EctA, EctC) and ectoine hydroxylase (EctD) for hydroxyectoine biosynthesis in the transcriptome of *Halocafeteria seosinensis*. Both studies established the possibility that the usage of the osmoprotective function of Ect is not limited to bacteria and archaea (Czech et al., 2018). Screening for potential enzymes involved in ectoine biosynthesis revealed that *S. salinarum* putatively possesses the complete functional pathway for the *de novo* synthesis of Ect containing the enzymes aspartate kinase (AsK), aspartate-semialdehyde dehydrogenase (AsD), diaminobutyrate-2-oxoglutarate aminotransferase (EctB), diaminobutyrate acetytransferase (EctA) and ectoine synthase (EctC; Suppl. Fig. 8). In contrast to Harding et al. (2016), EctA was the only enzyme showing a differential expression over time (Fig. 37). This might indicate together with the missing differential expression of the other enzymes that EctA could be the only occurring bottle neck for ectoine synthesis in *S. salinarum*. With this finding of potential Ect synthesis genes in the transcriptome of *S. salinarum*, evidence was provided for the possibility of a *de novo* synthesis of ectoine. An additional uptake of ectoine from the surrounding environment could not be proven as it was suggested before by Czech and Bremer (2018) in relation to the results of Weinisch et al. (2018b) because the former mentioned putative ectoine/proline and ectoine transporter were not up-regulated time-dependently. Concerning the maximum salinity tolerance of *S. salinarum* (up to 220 ‰), an increased ectoine import capacity would only be measurable either at higher salinities (> 90 ‰) or if choline and GB would be removed from the medium and replaced by additionally provided ectoine. The origin of the genes for ectoine synthesis and the localisation of their coded enzymes within the cell were discussed by Harding et al. (2017), Czech et al. (2018) and Czech and Bremer (2018) suggesting that ectoine synthesis in eukaryotes could happen in mitochondria using intermediates of the Krebs cycle and originate via lateral gene transfer by stealing genes from their food prokaryotes. To verify these assumptions for

S. salinarum, further detailed investigations including overexpression, isolation and functionality tests of the specific enzymes would be necessary.

Besides the former mentioned increased transport capacity for proline after 720 min, a significant and increasing up-regulation of potential enzymes for proline synthesis starting at 60 min and lasting until 720 min was measured (Fig. 37). Considering the conclusion that proline might be used as compatible solute by *S. salinarum* although proline was not detectable via ^{1}H-NMR before (Weinisch et al., 2018b), the transcriptome of *S. salinarum* presumably demonstrates a time-dependent order of compatible solute synthesis after salt up-shock, starting with choline and GB, respectively, followed by Ect and completed with proline. For instance, comparable behaviours were reported for some bacteria (*Halobacillus halophilus*, *Virgibacillus pantothenticus*) and other microorganisms, which produce several organic osmolytes, showing a temporal hierarchy of the dominantly synthesized compatible solute (Czech et al., 2018).

Results of ion imaging experiments of Weinisch et al. (2018b) revealed that *S. salinarum* is a 'low-salt – organic-solutes-in' strategist showing no significant cytoplasmic Na$^+$ or K$^+$ increase with increasing salinity. Conducted short-time adaptation experiments of single cells after a fast extracellular salinity increase demonstrated also no significant changes in cytoplasmic Na$^+$ concentrations (Weinisch et al., 2018b). To investigate how *S. salinarum* expulls additional ions from the cytoplasma and which transporters and channels are possibly involved chronologically after salt up-shock, thirteen different putative ion transporters and transport mechanisms were detected in the transcriptome of *S. salinarum* showing a time-dependent expression pattern (Fig. 38). Assuming that the predicted sub-cellular locations are correct and reliable for the nineteen different assumed ion transport capacities, it is considered that the annotated ion transport mechanisms played presumably a role in haloadaptation and were not masked by salinity-independent ion homeostasis mechanisms (Fig. 38). After a sudden increase of external salinity resulting in hyperosmotic conditions, a passively Na$^+$ influx together with a passive efflux of water molecules occurred. As an early response to this critical danger, putative plastid located mechanosensitive ion channels were significantly up-regulated after 2 and 10 min (Fig. 38). Mechanosensitive ion channels are known predominantly from bacteria as unspecific ion channels, which are gated by membrane tension and act as biosensors for

osmoregulation (Martinac and Kloda, 2003; Roberts, 2004). Even in archaea and eukaryotes, homologs of this channel family share the property of being activated by mechanical stimuli, play a role in turgor regulation (Martinac and Kloda, 2003) and releases mainly K^+ and Cl^- after hypoosmotic shock (Kloda and Martinac, 2001). Because of their non-directional ion release exclusively along the concentration gradient, it can be debated whether plastids in *S. salinarum* function as place for ion sequestration absorbing harmful Na^+ from the cytoplasm due to passive influx or as ion storage, where K^+ and Cl^- accumulate temporarily and then are released into the cytoplasm to balance the external salinity and to prevent water efflux. Another hint for an intracellular compartmentalization was possibly detected with the up-regulation of several putative channels presumably located in the endoplasmic reticulum and the lysosomes or vacuoles after 30 min (Fig. 38) transporting predominantly K^+, Cl^-, Ca^{2+} and anions because Grabe and Oster (2001) reported that cell organelles are more acidic in comparison to the cytoplasm concluding they are less sensitive to higher inorganic ion concentrations. Additionally, Muallem et al. (1985) reported that the endoplasmic reticulum (ER) function as Ca^{2+} storage and the amounts of Ca^{2+} in cytoplasm and organelles are different (Clapham, 2007). For these reasons, the mentioned higher transport capacity for K^+, Cl^- and Ca^{2+} after 30 min (Fig. 38) possibly illustrated that K^+ and Cl^- enter the cytoplasm to prevent an increased Na^+ influx (cf. Muallem et al., 1985) and Ca^{2+} is released for signalling (Clapham, 2007). Consequently, this potential ion flow together with the up-regulation of $Na^+/P_{inorganic}$ co-transporter, which might expulse Na^+ from the cytoplasm, explain the simultaneously down-regulation of K^+ transporting ATPases presumably located in the cell membrane because it seems that *S. salinarum* tried to compensate the higher external Na^+ concentration after salt up-shock first with an unspecific ion release from plastids. Then, if it was not effectual, the specific accumulation of K^+ and Cl^- originating from other cell compartments started after 30 min to prevent an increased Na^+ influx until the synthesis mechanisms for compatible solutes function sufficiently. Therefore, *S. salinarum* needed a decreased expulsion capacity of K^+ from the cytoplasm resulting in the mentioned down-regulation of putative K^+ transporting ATPases of the cell membrane at 30 min. Considering the former results implying that the compatible solute synthesis started after 30 min and lasted until 720 min after salt up-shock, the prevention of Na^+ influx by accumulating K^+ and Cl^- to compensate the osmotic pressure was not sufficient enough and only temporary restricted. Hence, the assumed capacity for active Na^+ expulsion from the

cytoplasm was increased after 60 and 120 min with the significant up-regulation of Na$^+$/H$^+$ antiporters presumably located in the cell membrane (Fig. 38). Due to the assumable slowly but constantly increasing amount of compatible solutes inside *S. salinarum* via synthesis, which might be demonstrated by the time-dependent up-regulation of the different required enzymes before, it could be assumed that the potentially accumulated ions in the cytoplasm after 2, 10 and 30 min have to be replaced by compatible solutes in order to not interfere cellular functions. Comparable observations were reported by Kempf and Bremer (1998) for the model organisms *Escherichia coli* and *Bacillus subtilis*, who accumulated K$^+$ via rapid uptake as an emergency reaction and replaced their initially amassed K$^+$ pool by compatible solutes (Roeßler and Müller, 2001; Czech et al., 2018). For this reason, putative mitochondrial Ca^{2+} transporting ATPases, Na$^+$/K$^+$ transporting ATPases and Na$^+$/H$^+$ antiporters of the endoplasmic reticulum as well as cation transporting ATPases of the cell membrane were significantly up-regulated after 720 min (Fig. 38). Combining the received results from the transcriptome of *S. salinarum*, the intracellular reaction to a salt up-shock could primarily consist of the following processes: After detection of an increased extracellular salinity, *S. salinarum* first releases K$^+$, Cl$^-$ and possibly Ca^{2+} from cell organelles into the cytoplasm to prevent Na$^+$ influx and later replaces them gradually with synthesized compatible solutes. During compatible solute synthesis, Na$^+$ ions are actively expulsed from the cytoplasm.

Besides the former mentioned possible haloadaptation processes and considering that a salt up-shock provokes not only osmotic stress, additional relevant metabolic functions in *S. salinarum* were investigated. As a consequence of cellular stress, the amount of reactive oxygen species (ROS) could increase resulting in a damage of proteins. Therefore, stress management of the cell could also be important except the basically osmoprotection (Harding et al., 2017). This was potentially observed in *S. salinarum* with the slightly up-regulation of putative enzymes involved in ROS detoxification and a huge amount of different heat shock proteins being mutually up-regulated at different time points (Fig. 39). Because of the lower oxygen solubility at higher salinities, the predominantly up-regulation of fermentative enzymes between 10 and 60 min after salt up-shock could be a cellular reaction to this physical effect. But the later occurring predominant up-regulation of putative enzymes for heme synthesis as well as for O$_2$ respiration (Fig. 39) exposed an intracellular oxygen deficiency as

unreliable. This conclusion is supported by the heterogenous regulation pattern of the TCA cycle with a constantly up-regulation of 2-methylisocitrate lyase (Fig. 39). Interestingly, the putative phosphoethanolamine-cytidyltransferase for phosphatidyl-ethanolamine synthesis was up-regulated after 30 min and the enzyme sterol-24-C-methyltransferase for possible sterol biosynthesis showed an up-regulation after 120 min (Fig. 39). These findings indicate that even the composition of cell membranes might be adjusted temporary under osmotic stress. Furthermore, the expression patterns of enzymes for ubiquinone synthesis (Fig. 39) could be connected directly with the higher energy demand of the cell caused by compatible solute synthesis because early after salt up-shock, no up-regulation was detected and between 30 and 120 min after salt up-shock the pronounced up-regulation occurred simultaneously with the above discussed compatible solute synthesis of mainly GB and Ect. In addition, the necessity of permanent active sensing processes might be illustrated with the significant up-regulation of the different phosphodiesterases belonging to cyclic nucleotide-dependent signalling at all time points except at 720 min (Fig. 39). If all these findings were compared to the reported differentially expressed genes as a function of salinity within *Halocafeteria seosinensis* (Harding and Simpson, 2018), significant and far-reaching differences are observed because Harding et al. (2017) investigated gene expression patterns of *Halocafeteria seosinensis* only in the steady state of the organism between two salinities and not in a time-resolved manner as this study does. Therefore, it is still unclear whether both organisms behave really differently according to their haloadaptation mechanisms or if the observed differences resulted potentially from different study designs.

To distinguish whether *S. salinarum* has the potential to tolerate higher amounts of inorganic ions in its cytoplasm, namely Na^+, or the former mentioned conclusion that *S. salinarum* uses the 'low-salt – organic-solutes-in' strategy is appropriate, the isoelectric points (pI) of the predicted proteins of *S. salinarum* were compared to isoelectric point profiles of proteins from the halophile bacterium *Salinibacter ruber*, the halophile black yeast *Hortaea wernickii* and the freshwater ciliate *Oxytricha trifallax*. Because these three organisms use different haloadaptation strategies, it can be assumed that if the transcriptome of *S. salinarum* reveals a similar proteome acidity pattern, *S. salinarum* might be able to use a similar haloadaptation strategy. Due to the majority of proteins (69.4 %) from *S. salinarum* showing more alkaline isoelectric points

(≥ 8) and less than 18.9 % of proteins could be connected to acidic values (≤ 6; Fig. 40), the proteome acidity of *S. salinarum* resemble the measured proteome acidity of *O. trifallax*, who is a close related species and as freshwater organism highly sensitive to increased intracellular ion concentrations. Investigating only the 15,093 ORFs of *S. salinarum* having at least a 4-fold-change once compared to t0, inconsequential higher proportion of proteins with more acidic values (≤ 6) could be observed (Fig. 40). This supports the former conclusions of both, this study and Weinisch et al. (2018b) that *S. salinarum* is a 'low-salt – organic-solutes-in' strategist because *S. ruber* exhibited as known 'high-salt-in' strategist significant more acidic proteins (Fig. 40). Considering the former argumentation of potential time-dependent ion flow within *S. salinarum* between certain cell compartments, differences according to the intracellular ion tolerances symbolized by the proteome acidity, should occur between different cell compartments of *S. salinarum*. As expected, the isoelectric point calculation of the 168,454 proteins from the transcriptome of *S. salinarum* provided the result that the relative amount of proteins sharing the same pI value varied obviously between cell compartments (Fig. 41). On the basis of this isoelectric point calculation, the pI values of the majority of proteins were taken as indicator for the proteome acidity of the cell compartment demonstrating large differences between cell compartments according to acidity (cf. Grabe and Oster, 2001). The lysosomes and vacuoles seem to be the most acidic organelles in *S. salinarum* similar to *Dunaliella salina* (Weiss and Pick, 1991). Therefore, it is most likely that vacuoles can be used for Na^+ expulsion from the cytoplasm supported by the detected putative vacuolar $Na^+/P_{inorganic}$ co-transporter and former observations of Na^+ entrapments at 210 ‰ by Weinisch et al. (2018b). In contrast, according to the distribution of the pI values of their proteins, mitochondria and plastids might be very sensitive to higher intracellular ion concentrations (Fig. 41). Especially for plastids, this conclusion conflicts with the former interpretation of the early up-regulated mechanosensitive ion channels (Fig. 38). Considering that the compatible solute synthesis might occur in mitochondria or plastids as it was reported for some plant species (Zhang et al., 2008), the proteins of mitochondria or plastids might not have to be acidic to tolerate higher ion concentrations because synthesized compatible solutes would balance them. Anonther possibility could be that even compatible solutes were additionally released into the cytoplasm via mechanosensitive channels. Furthermore, the majority of proteins located in the endoplasmic reticulum (ER) possessed only slightly alkaline pI values and thereby, they are more acidic than

proteins of plastids or mitochondria (Fig. 41). This could be a hint for a slightly higher ion tolerance in the ER than in mitochondria or plastids and therefore the ER could possibly function as ion repository (cf. Muallem et al., 1985; Koch, 1990). Interestingly, the majority of cytoplasmic proteins exhibited alkaline isoelectric points (44.5 % of proteins ≥ pI 8) but 34.9 % of all cytoplasmic proteins had more acidic pI values (≤ 6; Fig. 41). In conclusion, the cytoplasm of *S. salinarum* can only be named slightly acidic, if at all, due to the content of partly acidic proteins and therefore, *S. salinarum* might tolerate higher intracellular ion concentrations temporary as a short-time reaction to salt up-shock. Based on the organelle acidity, only lysosomes and vacuoles could be taken into account for the permanent accumulation of higher inorganic ion concentrations supporting the determination that *S. salinarum* uses the 'low-salt – organic-solutes-in' strategy.

In summary, after induced salt up-shock, *S. salinarum* presumably detected the increased salinity via two-component signal transduction systems and according to the HOG pathway, comparable osmosening mechanisms containing MAPK cascades. Although the predicted locations for the putative transporting mechanisms could be partly questioned, in an early reaction, K^+, Cl^- and also signalling relevant Ca^{2+} were possibly released from cell organelles into the cytoplasm to prevent Na^+ influx. Later, these inorganic ions should be replaced gradually by synthesized compatible solutes. The assumed compatible solute syntheses were performed in a time-dependent order, starting with choline repectively GB, followed by Ect and completed with proline. Additionally, choline and proline could be taken up via transporters by *S. salinarum*. During compatible solute accumulation, Na^+ ions were actively exported from the cytoplasm and ion compartmentalization between cell organelles might occur.

V.   Conclusions and Outlook

The emergence of high-salt habitats is progressing continuously, influenced either by anthropogenic actions or by climate change. Knowledge regarding ecological capacities and functions, which are based on the microbial inventory, i. e. their diversity and community structure, is however scarce. Nevertheless, salinity has already been identified as a strong environmental factor structuring communities, which makes the transition from e. g. freshwater to marine habitats for many organisms a challenge or even impossible. Within my PhD thesis, I could show that also microeukaryotes are subjected to specific salinity-dependent transition boundaries occurring along brackish to extreme hypersaline conditions. Using metatranscriptome analyses I aimed to link the observed protistan community shifts with distinct cellular haloadaptation strategies, as was hypothesized as explanation for the existence of salinity-dependent transition boundaries. The 'low-salt – organic-solutes-in' strategy appeared to be a preferred osmoadaptation strategy in hypersaline communities. The observation of a broad repertoire of gene regions presumably coding for different compatible solutes points to a potentially much longer list of compatible solutes known to be used by microeukaryotes, or protists in specific. These findings provide ground for further research to validate the metatranscriptome results, especially because most biosynthetic pathways are unknown in the field of protozoology. For example, this could be achieved using experimental approaches, such as quantitative real-time PCR (qRT-PCR) or the recently established ^{1}H-NMR spectroscopy for ciliates, to screen for and quantify compatible solutes in various protists isolated from different hypersaline environments. (Comparative) transcriptome data of various halotolerant and halophile protistan species, as well as the data of this study allow the extraction of gene information, which could be used to amplify genes of interest by PCR using genomic DNA of the species. Then, the obtained nucleotide fragments could be cloned into specific expression vectors, overexpressed and purified. Afterwards, proteins could be used for enzymatic analysis to reconstitute compatible solute pathways *in vitro* or for biotechnological applications, if possible.

Characterizing the 'low-salt – organic-solutes-in' strategy used by the halophile *Schmidingerothrix salinarum* in more detail, I was able to shed light on timed intracellular expression processes initiated by an external salinity increase. The observation of a possible serial synthesis of compatible solutes raises further questions,

such as what are the benefits of a serial synthesis or about the cell performance if genes for the synthesis of a specific compatible solute would be knocked-out.

With the results of my PhD thesis, I thus contributed to fill the gap of knowledge regarding the ecophysiology and distribution patterns of microeukaryotes thriving in hypersaline environments. Future research should focus therefore on proteome acidity as well as on direct measurements of intracellular ion concentrations of selected species to verify the predicted intracellular ion flows and to carify whether other haloadaptation strategies are also common in protists.

VI.   Summary/Zusammenfassung

Salinity is one of the most structuring environmental factors for microeukaryotic communities. Using eDNA barcoding, I detected significant shifts in microeukaryotic community compositions occurring at distinct salinities between brackish and marine conditions in the Baltic Sea. I, furthermore, conducted a metadata analysis including my and other marine and hypersaline community sequence data to confirm the existence of salinity-related transition boundaries and significant changes in alpha diversity patterns along a brackish to hypersaline gradient. One hypothesis for the formation of salinity-dependent transition boundaries between brackish to hypersaline conditions is the use of different cellular haloadaptation strategies. To test this hypothesis, I conducted metatranscriptome analyses of microeukaryotic communities along a pronounced salinity gradient (40 – 380 ‰). Clustering of functional transcripts revealed differences in metabolic properties and metabolic capacities between microeukaryotic communities at specific salinities, corresponding to the transition boundaries already observed in the taxonomic eDNA barcoding approach. In specific, microeukaryotic communities thriving at mid-hypersaline conditions (≤ 150 ‰) seem to predominantly apply the 'low-salt – organic-solutes-in' strategy by accumulating compatible solutes to counteract osmotic stress. Indications were found for both the intracellular synthesis of compatible solutes as well as for cellular transport systems. In contrast, communities of extreme-hypersaline habitats (≥ 200 ‰) may preferentially use the 'high-salt-in' strategy, i. e. the intracellular accumulation of inorganic ions in high concentrations, which is implied by the increased expression of Mg^{2+}, K^+, Cl^- transporters and channels.

In order to characterize the 'low-salt – organic-solutes-in' strategy applied by protists in more detail, I conducted a time-resolved transcriptome analysis of the heterotrophic ciliate *Schmidingerothrix salinarum* serving as model organism. *S. salinarum* was thus subjected to a salt-up shock to investigate the intracellular response to osmotic stress by shifts of gene expression. After increasing the external salinity, an increased expression of two-component signal transduction systems and MAPK cascades was observed. In an early reaction, the expression of transport mechanisms for K^+, Cl^- and Ca^{2+} increased, which may enhance the capacity of K^+, Cl^- and Ca^{2+} in the cytoplasm to compensate possibly harmful Na^+ influx. Expression of enzymes for the synthesis of possible compatible solutes, starting with glycine betaine, followed by ectoine and later proline, could imply that the inorganic ions K^+, Cl^- and Ca^{2+} are gradually replaced by the synthesized compatible solutes. Additionally, expressed transporters for choline (precursor of glycine betaine) and proline could indicate an intracellular accumulation of compatible solutes to balance the external salinity. During this accumulation, the up-regulated ion export mechanisms may increase the capacity for Na^+ expulsion from the cytoplasm and ion compartmentalization between cell organelles seem to happen.

The results of my PhD project revealed first evidence at molecular level for the salinity-dependent use of different haloadaptation strategies in microeukaryotes and significantly extend existing knowledge about haloadaptation processes in ciliates. The results provide ground for future research, such as (comparative) transcriptome analysis of ciliates thriving in extreme-hypersaline habitats or experiments like qRT-PCR to validate transcriptome results.

Der Salzgehalt ist einer der am stärksten strukturierenden Umweltfaktoren für mikroeukaryotische Lebensgemeinschaften. Mittels eDNA-Barcoding konnte ich in der Ostsee signifikante Verschiebungen in den mikroeukaryotischen Gemeinschaftszusammensetzungen bei unterschiedlichen Salinitäten zwischen brackigen und marinen Bedingungen aufdecken. Außerdem führte ich eine Metadatenanalyse durch, die neben meinen auch andere Sequenzdaten aus marinen und hypersalinen Gemeinschaften beinhaltete, um die Existenz von salinitätsbedingten Übergangsgrenzen und signifikanten Veränderungen in den Alpha-Diversitätsmustern entlang eines brackigen bis hypersalinen Gradienten zu bestätigen. Eine mögliche Erklärung für die salinitätsabhängigen Übergangsgrenzen ist die Verwendung verschiedener zellulärer Salzanpassungsstrategien. Um diese Hypothese zu testen, habe ich Metatranskriptomanalysen von mikroeukaryotischen Gemeinschaften entlang eines ausgeprägten Salinitätsgradienten (40 - 380 ‰) durchgeführt. Die Gruppierung funktioneller Transkripte zeigte Unterschiede in den metabolischen Eigenschaften und Stoffwechselkapazitäten zwischen mikroeukaryotischen Gemeinschaften aus bestimmten Salinitäten, die den Übergangsgrenzen aus dem taxonomischen eDNA-Barcoding entsprechen. Insbesondere mikroeukaryotische Gemeinschaften aus mittel-hypersalinen Bedingungen (≤ 150 ‰) scheinen dem osmotischem Stress entgegenzuwirken in dem sie überwiegend die "Low-Salt - Organic-Solutes-in" Strategie durch Anreicherung kompatibler Solute verwenden. Hinweise wurden sowohl für die intrazelluläre Synthese von kompatiblen Soluten als auch für zelluläre Transportsysteme gefunden. Im Gegensatz dazu scheinen Gemeinschaften aus extrem-hypersalinen Lebensräumen (≥ 200 ‰) eher die "High-Salt-in" Strategie zu nutzen (intrazelluläre Anreicherung anorganischer Ionen in hohen Konzentrationen), was durch die erhöhte Expression von Mg^{2+}, K^+, Cl^- -Ionentransporter und -Kanäle impliziert wird.

Zu einer genaueren Charakterisierung der von Protisten genutzten "Low-Salt - Organic-Solutes-in" Strategie, habe ich eine zeitaufgelöste Transkriptomanalyse des heterotrophen Ciliaten *Schmidingerothrix salinarum* durchgeführt. Hierzu wurde *S. salinarum* einem Salzschock ausgesetzt und anschließend anhand Gen-Expressionsveränderungen die intrazelluläre Reaktion auf osmotischen Stress untersucht. Nach Erhöhung des externen Salzgehalts wurde eine erhöhte Expression von Zweikomponenten-Signaltransduktionssysteme und MAPK-Kaskaden festgestellt. Zuerst werden vermehrt Transportmechanismen für K^+, Cl^- und Ca^{2+}-Ionen exprimiert, die eventuell für eine erhöhte Kapazität von K^+, Cl^- und Ca^{2+}-Ionen im Zytoplasma sorgen um möglicherweise einem schädlichen Na^+-Einstrom entgegenzuwirken. Die nachfolgende Enzymexpression für die Synthese möglicher kompatibler Solute, beginnend mit Glycinbetain, gefolgt von Ectoin und später Prolin, könnte implizieren, dass die anorganischen Ionen K^+, Cl^- und Ca^{2+} im Zytoplasma schrittweise durch synthetisierten kompatiblen Solute ersetzt werden. Zusätzlich exprimierte Transporter für Cholin (Vorläufer von Glycinbetain) und Prolin könnten neben der Synthese für eine intrazelluläre Akkumulierung von kompatiblen Soluten sprechen, um den externen Salzgehalt auszugleichen. Während dieser Akkumulation könnten die hochregulierten Ionen-Exportmechanismen für eine erhöhte Na^+-Ausschlusskapazität aus dem Zytoplasma sorgen, wodurch sich eine Ionen-Kompartimentierung zwischen den Zellorganellen andeuten würde.

Die Ergebnisse meiner Dissertation lieferten erstmalig Hinweise auf molekularer Ebene für die salzabhängige Nutzung verschiedener Haloadaptionsstrategien in Mikroeukaryoten und erweiterten das vorhandene Wissen über Haloadaptationsprozesse bei Ciliaten erheblich. Die Ergebnisse bilden die Grundlage für zukünftige Forschungen, wie z. B. Experimente (qRT-PCR) zur Validierung von Transkriptomergebnissen oder (vergleichende) Transkriptomanalysen von Ciliaten, die in extrem-hypersalinen Lebensräumen gedeihen.

VII. References

Aeschlimann, S.H., Jönsson, F., Postberg, J., Stover, N.A., Petera, R.L., Lipps, H.-J., Nowacki, M., Swart, E.C., 2014. The draft assembly of the radically organized Stylonychia lemnae macronuclear genome. Genome biology and evolution 6 (7), 1707–1723. 10.1093/gbe/evu139.

Almagro Armenteros, J.J., Sønderby, C.K., Sønderby, S.K., Nielsen, H., Winther, O., 2017. DeepLoc: Prediction of protein subcellular localization using deep learning. Bioinformatics (Oxford, England) 33 (21), 3387–3395. 10.1093/bioinformatics/btx431.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. Journal of Molecular Biology 215 (3), 403–410. 10.1016/S0022-2836(05)80360-2.

Andrews, S., 2012. FastQC a quality control tool for high throughput sequence data.

Anonymous, 1958. The Venice System for the Classification of Marine Waters According to Salinity. Limnol. Oceanogr. 3 (3), 346–347. 10.4319/lo.1958.3.3.0346.

Ariño, J., Ramos, J., Sychrová, H., 2010. Alkali metal cation transport and homeostasis in yeasts. Microbiology and molecular biology reviews : MMBR 74 (1), 95–120. 10.1128/MMBR.00042-09.

Atkinson, M.J., Bingman, C., 1997. Elemental composition of commercial seasalts. Journal of Aquariculture and Aquatic Sciences 8 (2), 39.

Aury, J.-M., Jaillon, O., Duret, L., Noel, B., Jubin, C., Porcel, B.M., Ségurens, B., Daubin, V., Anthouard, V., Aiach, N., Arnaiz, O., Billaut, A., Beisson, J., Blanc, I., Bouhouche, K., Câmara, F., Duharcourt, S., Guigo, R., Gogendeau, D., Katinka, M., Keller, A.-M., Kissmehl, R., Klotz, C., Koll, F., Le Mouël, A., Lepère, G., Malinsky, S., Nowacki, M., Nowak, J.K., Plattner, H., Poulain, J., Ruiz, F., Serrano, V., Zagulski, M., Dessen, P., Bétermier, M., Weissenbach, J., Scarpelli, C., Schächter, V., Sperling, L., Meyer, E., Cohen, J., Wincker, P., 2006. Global trends of whole-genome duplications revealed by the ciliate Paramecium tetraurelia. Nature 444 (7116), 171–178. 10.1038/nature05230.

Bachy, C., Dolan, J.R., López-García, P., Deschamps, P., Moreira, D., 2013. Accuracy of protist diversity assessments: Morphology compared with cloning and direct pyrosequencing of 18S rRNA genes and ITS regions using the conspicuous tintinnid ciliates as a case study. The ISME journal 7 (2), 244–255. 10.1038/ismej.2012.106.

Baldauf, S.L., 2008. An overview of the phylogeny and diversity of eukaryotes. Journal of Systematics and Evolution 46 (3), 263–273.

Barber, R.T., 2007. Oceans. Picoplankton do some heavy lifting. Science (New York, N.Y.) 315 (5813), 777–778. 10.1126/science.1137438.

Beisser, D., Graupner, N., Bock, C., Wodniok, S., Grossmann, L., Vos, M., Sures, B., Rahmann, S., Boenigk, J., 2017. Comprehensive transcriptome analysis provides new insights into nutritional strategies and phylogenetic relationships of chrysophytes. PeerJ 5, e2832. 10.7717/peerj.2832.

Ben-Amotz, A., Avron, M., 1980. Osmoregulation in the halophilic algae Dunaliella and Asteromonas, in: , Genetic engineering of osmoregulation. Springer, pp. 91–99.

Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Wheeler, D.L., 2005. GenBank. Nucleic acids research 33 (Database issue), D34-8. 10.1093/nar/gki063.

Benton, M.J., 2009. The Red Queen and the Court Jester: Species diversity and the role of biotic and abiotic factors through time. Science (New York, N.Y.) 323 (5915), 728–732.

Bickel-Sandkötter, S., Ufer, M., Steinert, K., Dane, M., 1995. Leben im Salz — Halophile Archaea. Biologie in unserer Zeit 25 (6), 380–389. 10.1002/biuz.19950250624.

Bills, B.G., Borsa, A.A., Comstock, R.L., 2007. MISR-based passive optical bathymetry from orbit with few-cm level of accuracy on the Salar de Uyuni, Bolivia. Remote Sensing of Environment 107 (1-2), 240–255. 10.1016/j.rse.2006.11.006.

Blumwald, E., Mehlhorn, R.J., Packer, L., 1983. Ionic Osmoregulation during Salt Adaptation of the Cyanobacterium Synechococcus 6311. Plant physiology 73 (2), 377–380. 10.1104/pp.73.2.377.

Boch, J., Kempf, B., Schmid, R., Bremer, E., 1996. Synthesis of the osmoprotectant glycine betaine in Bacillus subtilis: Characterization of the gbsAB genes. J. Bacteriol. 178 (17), 5121–5129. 10.1128/jb.178.17.5121-5129.1996.

Boenigk, J., Wodniok, S., Bock, C., Beisser, D., Hempel, C., Grossmann, L., Lange, A., Jensen, M., 2018. Geographic distance and mountain ranges structure freshwater protist communities on a European scale. Metabarcoding and Metagenomics 2, e21519. 10.3897/mbmg.2.21519.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics (Oxford, England) 30 (15), 2114–2120. 10.1093/bioinformatics/btu170.

Bornette, G., Puijalon, S., 2011. Response of aquatic plants to abiotic factors: A review. Aquat Sci 73 (1), 1–14. 10.1007/s00027-010-0162-7.

Brown, A.D., 1976. Microbial Water Stress. Bacteriological Reviews 40 (4), 803–846.

Brown, A.D., 1990. Microbial water stress physiology: Principles and perspectives. Wiley, Chichester, New York, xiii, 313.

Butschinsky, P., 1897. Die Protozooen-Fauna der Salzsee-Limane bei Odessa. Zoologischer Anzeiger 20 (533), 194–197.

Campbell, B.J., Kirchman, D.L., 2013. Bacterial diversity, community structure and potential growth rates along an estuarine salinity gradient. The ISME journal 7 (1), 210–220. 10.1038/ismej.2012.93.

Cánovas, D., Vargas, C., Kneip, S., Morón, M.J., Ventosa, A., Bremer, E., Nieto, J.J., 2000. Genes for the synthesis of the osmoprotectant glycine betaine from choline in the moderately halophilic bacterium Halomonas elongata DSM 3043, USA. Microbiology (Reading, England) 146 (Pt 2), 455–463. 10.1099/00221287-146-2-455.

Cantrell, S.A., Casillas-Martínez, L., Molina, M., 2006. Characterization of fungi from hypersaline environments of solar salterns using morphological and molecular techniques. Mycological research 110 (Pt 8), 962–970. 10.1016/j.mycres.2006.06.005.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., Knight, R., 2010. QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7, 335 EP -. 10.1038/nmeth.f.303.

Capra, E.J., Laub, M.T., 2012. Evolution of two-component signal transduction systems. Annual review of microbiology 66, 325–347. 10.1146/annurev-micro-092611-150039.

Cavalcanti, A.R.O., Stover, N.A., Orecchia, L., Doak, T.G., Landweber, L.F., 2004. Coding properties of Oxytricha trifallax (Sterkiella histriomuscorum) macronuclear chromosomes: Analysis of a pilot genome project. Chromosoma 113 (2), 69–76. 10.1007/s00412-004-0295-3.

Chapin, F.S., Walker, B.H., Hobbs, R.J., Hooper, D.U., Lawton, J.H., Sala, O.E., Tilman, D., Chapin III, F.S., 1997. Biotic Control over the Functioning of Ecosystems. Science 277 (5325), 500–504. 10.1126/science.277.5325.500.

Chen, Y., Lun, A.T.L., Smyth, G.K., op. 2014. Differential Expression Analysis of Complex RNA-seq Experiments Using edgeR, in: Datta, S., Nettleton, D. (Eds.), Statistical analysis of next generation sequencing data. Springer, Cham, pp. 51–74.

Christian, J.H.B., Waltho, J.A., 1962. Solute concentrations within cells of halophilic and non-halophilic bacteria. Biochimica et Biophysica Acta 65 (3), 506–508. 10.1016/0006-3002(62)90453-5.

Clapham, D.E., 2007. Calcium signaling. Cell 131 (6), 1047–1058. 10.1016/j.cell.2007.11.028.

Crump, B.C., Hopkinson, C.S., Sogin, M.L., Hobbie, J.E., 2004. Microbial Biogeography along an Estuarine Salinity Gradient: Combined Influences of Bacterial Growth and Residence

Time. Applied and environmental microbiology 70 (3), 1494–1505. 10.1128/AEM.70.3.1494-1505.2004.

Czech, L., Bremer, E., 2018. With a pinch of extra salt-Did predatory protists steal genes from their food? PLoS biology 16 (2), e2005163. 10.1371/journal.pbio.2005163.

Czech, L., Hermann, L., Stöveken, N., Richter, A.A., Höppner, A., Smits, S.H.J., Heider, J., Bremer, E., 2018. Role of the Extremolytes Ectoine and Hydroxyectoine as Stress Protectants and Nutrients: Genetics, Phylogenomics, Biochemistry, and Structural Analysis. Genes 9 (4). 10.3390/genes9040177.

Danson, M.J., Hough, D.W., 1997. The Structural Basis of Protein Halophilicity. Comparative Biochemistry and Physiology Part A: Physiology 117 (3), 307–312. 10.1016/S0300-9629(96)00268-X.

DasSarma, S., Arora, P., 2002. Halophiles, encyclopedia of life sciences. 8: 458–466. London: Nature Publishing Group.

Dillies, M.-A., Rau, A., Aubert, J., Hennequet-Antier, C., Jeanmougin, M., Servant, N., Keime, C., Marot, G., Castel, D., Estelle, J., Guernec, G., Jagla, B., Jouneau, L., Laloë, D., Le Gall, C., Schaëffer, B., Le Crom, S., Guedj, M., Jaffrézic, F., 2013. A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. Briefings in bioinformatics 14 (6), 671–683. 10.1093/bib/bbs046.

Doak, T.G., Cavalcanti, A.R.O., Stover, N.A., Dunn, D.M., Weiss, R., Herrick, G., Landweber, L.F., 2003. Sequencing the Oxytricha trifallax macronuclear genome: A pilot project. Trends in genetics : TIG 19 (11), 603–607. 10.1016/j.tig.2003.09.013.

Dunson, W.A., Travis, J., 1991. The Role of Abiotic Factors in Community Organization. The American Naturalist 138 (5), 1067–1091. 10.1086/285270.

Dunthorn, M., Stoeck, T., Clamp, J., Warren, A., Mahé, F., 2014. Ciliates and the rare biosphere: A review. The Journal of eukaryotic microbiology 61 (4), 404–409. 10.1111/jeu.12121.

Dupont, C.L., Larsson, J., Yooseph, S., Ininbergs, K., Goll, J., Asplund-Samuelsson, J., McCrow, J.P., Celepli, N., Allen, L.Z., Ekman, M., Lucas, A.J., Hagström, Å., Thiagarajan, M., Brindefalk, B., Richter, A.R., Andersson, A.F., Tenney, A., Lundin, D., Tovchigrechko, A., Nylander, J.A.A., Brami, D., Badger, J.H., Allen, A.E., Rusch, D.B., Hoffman, J., Norrby, E., Friedman, R., Pinhassi, J., Venter, J.C., Bergman, B., 2014. Functional tradeoffs underpin salinity-driven divergence in microbial community composition. PloS one 9 (2), e89549. 10.1371/journal.pone.0089549.

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R., 2011. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics (Oxford, England) 27 (16), 2194–2200. 10.1093/bioinformatics/btr381.

Eisen, J.A., Coyne, R.S., Wu, M., Wu, D., Thiagarajan, M., Wortman, J.R., Badger, J.H., Ren, Q., Amedeo, P., Jones, K.M., Tallon, L.J., Delcher, A.L., Salzberg, S.L., Silva, J.C., Haas, B.J., Majoros, W.H., Farzad, M., Carlton, J.M., Smith, R.K., Garg, J., Pearlman, R.E., Karrer, K.M., Sun, L., Manning, G., Elde, N.C., Turkewitz, A.P., Asai, D.J., Wilkes, D.E., Wang, Y., Cai, H., Collins, K., Stewart, B.A., Lee, S.R., Wilamowska, K., Weinberg, Z., Ruzzo, W.L., Wloga, D., Gaertig, J., Frankel, J., Tsao, C.-C., Gorovsky, M.A., Keeling, P.J., Waller, R.F., Patron, N.J., Cherry, J.M., Stover, N.A., Krieger, C.J., del Toro, C., Ryder, H.F., Williamson, S.C., Barbeau, R.A., Hamilton, E.P., Orias, E., 2006. Macronuclear genome sequence of the ciliate Tetrahymena thermophila, a model eukaryote. PLoS biology 4 (9), e286. 10.1371/journal.pbio.0040286.

Elloumi, J., Carrias, J.-F., Ayadi, H., Sime-Ngando, T., Bouaïn, A., 2009. Communities structure of the planktonic halophiles in the solar saltern of Sfax, Tunisia. Estuarine, Coastal and Shelf Science 81 (1), 19–26. 10.1016/j.ecss.2008.09.019.

Empadinhas, N., da Costa, M.S., 2008. Osmoadaptation mechanisms in prokaryotes: distribution of compatible solutes. Int Microbiol 11 (3), 151–161. 10.2436/20.1501.01.55.

Epstein, S., López-García, P., 2008. "Missing" protists: A molecular prospective. Biodivers Conserv 17 (2), 261–276. 10.1007/s10531-007-9250-y.

Fan, F., Ghanem, M., Gadda, G., 2004. Cloning, sequence analysis, and purification of choline oxidase from Arthrobacter globiformis: A bacterial enzyme involved in osmotic stress tolerance. Archives of Biochemistry and Biophysics 421 (1), 149–158. 10.1016/j.abb.2003.10.003.

Filker, S., Forster, D., Weinisch, L., Mora-Ruiz, M., González, B., Farías, M.E., Rosselló-Móra, R., Stoeck, T., 2017. Transition boundaries for protistan species turnover in hypersaline waters of different biogeographic regions. Environmental microbiology 19 (8), 3186–3200. 10.1111/1462-2920.13805.

Filker, S., Gimmler, A., Dunthorn, M., Mahé, F., Stoeck, T., 2015. Deep sequencing uncovers protistan plankton diversity in the Portuguese Ria Formosa solar saltern ponds. Extremophiles : life under extreme conditions 19 (2), 283–295. 10.1007/s00792-014-0713-2.

Filker, S., Kühner, S., Heckwolf, M., Dierking, J., Stoeck, T., 2019. A fundamental difference between macrobiota and microbial eukaryotes: Protistan plankton has a species

maximum in the freshwater-marine transition zone of the Baltic Sea. Environmental microbiology 21 (2), 603–617. 10.1111/1462-2920.14502.

Foissner, W., Filker, S., Stoeck, T., 2014. Schmidingerothrix salinarum nov. spec. is the molecular sister of the large oxytrichid clade (ciliophora, hypotricha). The Journal of eukaryotic microbiology 61 (1), 61–74. 10.1111/jeu.12087.

Forster, D., Behnke, A., Stoeck, T., 2012. Meta-analyses of environmental sequence data identify anoxia and salinity as parameters shaping ciliate communities. Systematics and Biodiversity 10 (3), 277–288. 10.1080/14772000.2012.706239.

Forster, D., Dunthorn, M., Stoeck, T., Mahé, F., 2016. Comparison of three clustering approaches for detecting novel environmental microbial diversity. PeerJ 4, e1692. 10.7717/peerj.1692.

Fu, L., Niu, B., Zhu, Z., Wu, S., Li, W., 2012. CD-HIT: Accelerated for clustering the next-generation sequencing data. Bioinformatics (Oxford, England) 28 (23), 3150–3152. 10.1093/bioinformatics/bts565.

Galinski, E.A., Trüper, H.G., 1982. Betaine, a compatible solute in the extremely halophilic phototrophic bacterium Ectothiorhodospira halochloris. FEMS Microbiology Letters 13 (4), 357–360. 10.1111/j.1574-6968.1982.tb08287.x.

Galinski, E.A., Trüper, H.G., 1994. Microbial behaviour in salt-stressed ecosystems. FEMS Microbiology Reviews 15 (2-3), 95–108. 10.1111/j.1574-6976.1994.tb00128.x.

Gifford, S.M., Sharma, S., Rinta-Kanto, J.M., Moran, M.A., 2011. Quantitative analysis of a deeply sequenced marine microbial metatranscriptome. The ISME journal 5 (3), 461–472. 10.1038/ismej.2010.141.

Gilbert, J.A., Field, D., Huang, Y., Edwards, R., Li, W., Gilna, P., Joint, I., 2008. Detection of large numbers of novel sequences in the metatranscriptomes of complex marine microbial communities. PloS one 3 (8), e3042. 10.1371/journal.pone.0003042.

Gillaspy, G.E., 2011. The cellular language of myo-inositol signaling. The New phytologist 192 (4), 823–839. 10.1111/j.1469-8137.2011.03939.x.

Ginzburg, M., Sachs, L., Ginzburg, B.Z., 1970. Ion metabolism in a Halobacterium. I. Influence of age of culture on intracellular concentrations. The Journal of general physiology 55 (2), 187–207.

Giovannoni, S.J., Stingl, U., 2005. Molecular diversity and ecology of microbial plankton. Nature 437 (7057), 343–348. 10.1038/nature04158.

Go, Y.-M., Jones, D.P., 2013. Thiol/disulfide redox states in signaling and sensing. Critical reviews in biochemistry and molecular biology 48 (2), 173–181. 10.3109/10409238.2013.764840.

Gostincar, C., Turk, M., Plemenitas, A., Gunde-Cimerman, N., 2009. The expressions of Delta 9-, Delta 12-desaturases and an elongase by the extremely halotolerant black yeast Hortaea werneckii are salt dependent. FEMS yeast research 9 (2), 247–256. 10.1111/j.1567-1364.2009.00481.x.

Grabe, M., Oster, G., 2001. Regulation of Organelle Acidity. The Journal of general physiology 117 (4), 329–344. 10.1085/jgp.117.4.329.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature biotechnology 29 (7), 644–652. 10.1038/nbt.1883.

Gray, M.W., 2002. Mitochondria, in: , Encyclopedia of genetics. Acad. Press, San Diego, pp. 1215–1217.

Grossmann, L., Jensen, M., Heider, D., Jost, S., Glücksman, E., Hartikainen, H., Mahamdallie, S.S., Gardner, M., Hoffmann, D., Bass, D., Boenigk, J., 2016. Protistan community analysis: Key findings of a large-scale molecular sampling. The ISME journal 10 (9), 2269–2279. 10.1038/ismej.2016.10.

Gu, Z.J., Wang, L., Le Rudulier, D., Zhang, B., Yang, S.S., 2008. Characterization of the glycine betaine biosynthetic genes in the moderately halophilic bacterium Halobacillus dabanensis D-8(T). Current microbiology 57 (4), 306–311. 10.1007/s00284-008-9194-9.

Gunde-Cimerman, N., Plemenitaš, A., 2006. Ecology and molecular adaptations of the halophilic black yeast Hortaea werneckii. Rev Environ Sci Biotechnol 5 (2-3), 323–331. 10.1007/s11157-006-9105-0.

Gunde-Cimerman, N., Plemenitaš, A., Oren, A., 2018. Strategies of adaptation of microorganisms of the three domains of life to high salt concentrations. FEMS Microbiology Reviews 42 (3), 353–375. 10.1093/femsre/fuy009.

Gunter, G., 1961. Some relations of estuarine organisms to salinity. Limnology and Oceanography 6 (2), 182–190.

Gutowska, M.A., Melzner, F., 2009. Abiotic conditions in cephalopod (Sepia officinalis) eggs: Embryonic development at low pH and high pCO2. Mar Biol 156 (3), 515–519. 10.1007/s00227-008-1096-7.

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., Henschel, R., LeDuc, R.D., Friedman, 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. 10.1038/nprot.2013.084.

Hällfors, G., 2004. Checklist of Baltic Sea Phytoplankton species (including some Heterotrophic Protistan groups). Baltic Sea Environment Proceedings 95, 1–208.

Hänelt, I., Müller, V., 2013. Molecular Mechanisms of Adaptation of the Moderately Halophilic Bacterium Halobacillis halophilus to Its Environment. Life (Basel, Switzerland) 3 (1), 234–243. 10.3390/life3010234.

Harding, T., Brown, M.W., Simpson, A.G.B., Roger, A.J., 2016. Osmoadaptative Strategy and Its Molecular Signature in Obligately Halophilic Heterotrophic Protists. Genome biology and evolution 8 (7), 2241–2258. 10.1093/gbe/evw152.

Harding, T., Roger, A.J., Simpson, A.G.B., 2017. Adaptations to High Salt in a Halophilic Protist: Differential Expression and Gene Acquisitions through Duplications and Gene Transfers. Frontiers in microbiology 8, 944. 10.3389/fmicb.2017.00944.

Harding, T., Simpson, A.G.B., 2018. Recent Advances in Halophilic Protozoa Research. The Journal of eukaryotic microbiology 65 (4), 556–570. 10.1111/jeu.12495.

Hauer, G., Rogerson, A., 2005. Heterotrophic Protozoa from Hypersaline Environments, in: Adaptation to Life at High Salt Concentrations in Archaea, Bacteria, and Eukarya. Springer Netherlands, Dordrecht, pp. 519–539.

Heaphy, S.M., Mariotti, M., Gladyshev, V.N., Atkins, J.F., Baranov, P.V., 2016. Novel Ciliate Genetic Code Variants Including the Reassignment of All Three Stop Codons to Sense Codons in Condylostoma magnum. Molecular biology and evolution 33 (11), 2885–2889. 10.1093/molbev/msw166.

Heidelberg, K.B., Nelson, W.C., Holm, J.B., Eisenkolb, N., Andrade, K., Emerson, J.B., 2013. Characterization of eukaryotic microbial diversity in hypersaline Lake Tyrrell, Australia. Frontiers in microbiology 4, 115. 10.3389/fmicb.2013.00115.

Herlemann, D.P., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J.J., Andersson, A.F., 2011. Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. The ISME journal 5 (10), 1571–1579. 10.1038/ismej.2011.41.

Herrick, G., 1994. Germline-soma relationships in ciliated protozoa: The inception and evolution of nuclear dimorphism in one-celled animals. Seminars in Developmental Biology 5 (1), 3–12. 10.1006/sedb.1994.1002.

Hersen, P., McClean, M.N., Mahadevan, L., Ramanathan, S., 2008. Signal processing by the HOG MAP kinase pathway. Proceedings of the National Academy of Sciences of the United States of America 105 (20), 7165–7170. 10.1073/pnas.0710770105.

Hsieh, T.C., Ma, K.H., Chao, A., McInerny, G., 2016. iNEXT: An R package for rarefaction and extrapolation of species diversity (Hill numbers). Methods Ecol Evol 7 (12), 1451–1456. 10.1111/2041-210X.12613.

Hu, S.K., Liu, Z., Alexander, H., Campbell, V., Connell, P.E., Dyhrman, S.T., Heidelberg, K.B., Caron, D.A., 2018. Shifting metabolic priorities among key protistan taxa within and below the euphotic zone. Environmental microbiology 20 (8), 2865–2879. 10.1111/1462-2920.14259.

Hu, S.K., Liu, Z., Lie, A.A.Y., Countway, P.D., Kim, D.Y., Jones, A.C., Gast, R.J., Cary, S.C., Sherr, E.B., Sherr, B.F., Caron, D.A., 2015. Estimating Protistan Diversity Using High-Throughput Sequencing. The Journal of eukaryotic microbiology 62 (5), 688–693. 10.1111/jeu.12217.

Hu, Y.O.O., Karlson, B., Charvet, S., Andersson, A.F., 2016. Diversity of Pico- to Mesoplankton along the 2000 km Salinity Gradient of the Baltic Sea. Frontiers in microbiology 7, 679. 10.3389/fmicb.2016.00679.

Huang, Y., Niu, B., Gao, Y., Fu, L., Li, W., 2010. CD-HIT Suite: A web server for clustering and comparing biological sequences. Bioinformatics (Oxford, England) 26 (5), 680–682. 10.1093/bioinformatics/btq003.

Hugerth, L.W., Larsson, J., Alneberg, J., Lindh, M.V., Legrand, C., Pinhassi, J., Andersson, A.F., 2015. Metagenome-assembled genomes uncover a global brackish microbiome. Genome biology 16, 279. 10.1186/s13059-015-0834-7.

Imhoff, J.F., 1986. Osmoregulation and compatible solutes in eubacteria. FEMS Microbiology Letters 39 (1-2), 57–66. 10.1111/j.1574-6968.1986.tb01843.x.

IPCC, 2013. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, in: Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M.M.B., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M. (Eds.), Climate Change 2013. The physical science basis. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.

Jahn, C.L., Klobutcher, L.A., 2002. Genome remodeling in ciliated protozoa. Annual review of microbiology 56, 489–520. 10.1146/annurev.micro.56.012302.160916.

Javor, B., 1989. Hypersaline Environments: Microbiology and Biogeochemistry. Springer Berlin Heidelberg, Berlin, Heidelberg, 1 online resource (viii, 328.

Joghee, N.N., Jayaraman, G., 2014. Metabolomic characterization of halophilic bacterial isolates reveals strains synthesizing rare diaminoacids under salt stress. Biochimie 102, 102–111. 10.1016/j.biochi.2014.02.015.

Jost, L., 2007. Partitioning Diversity into independent Alpha and Beta Components. Ecology 88 (10), 2427–2439. 10.1890/06-1736.1.

Kanehisa, M., 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic acids research 28 (1), 27–30. 10.1093/nar/28.1.27.

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., Morishima, K., 2017. KEGG: New perspectives on genomes, pathways, diseases and drugs. Nucleic acids research 45 (D1), D353-D361. 10.1093/nar/gkw1092.

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., Tanabe, M., 2016a. KEGG as a reference resource for gene and protein annotation. Nucleic acids research 44 (D1), D457-62. 10.1093/nar/gkv1070.

Kanehisa, M., Sato, Y., Morishima, K., 2016b. BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. Journal of Molecular Biology 428 (4), 726–731. 10.1016/j.jmb.2015.11.006.

Keeling, P.J., Burki, F., Wilcox, H.M., Allam, B., Allen, E.E., Amaral-Zettler, L.A., Armbrust, E.V., Archibald, J.M., Bharti, A.K., Bell, C.J., Beszteri, B., Bidle, K.D., Cameron, C.T., Campbell, L., Caron, D.A., Cattolico, R.A., Collier, J.L., Coyne, K., Davy, S.K., Deschamps, P., Dyhrman, S.T., Edvardsen, B., Gates, R.D., Gobler, C.J., Greenwood, S.J., Guida, S.M., Jacobi, J.L., Jakobsen, K.S., James, E.R., Jenkins, B., John, U., Johnson, M.D., Juhl, A.R., Kamp, A., Katz, L.A., Kiene, R., Kudryavtsev, A., Leander, B.S., Lin, S., Lovejoy, C., Lynn, D., Marchetti, A., McManus, G., Nedelcu, A.M., Menden-Deuer, S., Miceli, C., Mock, T., Montresor, M., Moran, M.A., Murray, S., Nadathur, G., Nagai, S., Ngam, P.B., Palenik, B., Pawlowski, J., Petroni, G., Piganeau, G., Posewitz, M.C., Rengefors, K., Romano, G., Rumpho, M.E., Rynearson, T., Schilling, K.B., Schroeder, D.C., Simpson, A.G.B., Slamovits, C.H., Smith, D.R., Smith, G.J., Smith, S.R., Sosik, H.M., Stief, P., Theriot, E., Twary, S.N., Umale, P.E., Vaulot, D., Wawrik, B., Wheeler, G.L., Wilson, W.H., Xu, Y., Zingone, A., Worden, A.Z., 2014. The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. PLoS biology 12 (6), e1001889. 10.1371/journal.pbio.1001889.

Kellogg, C.A., Griffin, D.W., 2006. Aerobiology and the global transport of desert dust. Trends in ecology & evolution 21 (11), 638–644. 10.1016/j.tree.2006.07.004.

Kempf, B., Bremer, E., 1998. Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. Archives of Microbiology 170 (5), 319–330. 10.1007/s002030050649.

Kimor, B., Moigis, A.G., Dohms, V., Stienen, C., 1985. A case of mass occurrence of Prorocentrum minimum in the Kiel Fjord. Mar. Ecol. Prog. Ser 27, 209–215.

Kirby, H., 1932. Two protozoa from brine. Trans. Am. Microscop. (51), 8–15.

Klee, C.B., Crouch, T.H., Richman, P.G., 1980. Calmodulin. Annual review of biochemistry 49 (1), 489–515.

Kloda, A., Martinac, B., 2001. Structural and functional differences between two homologous mechanosensitive channels of Methanococcus jannaschii. The EMBO journal 20 (8), 1888–1896. 10.1093/emboj/20.8.1888.

Koch, G.L., 1990. The endoplasmic reticulum and calcium storage. BioEssays : news and reviews in molecular, cellular and developmental biology 12 (11), 527–531. 10.1002/bies.950121105.

Kogej, T., Ramos, J., Plemenitas, A., Gunde-Cimerman, N., 2005. The Halophilic Fungus Hortaea werneckii and the Halotolerant Fungus Aureobasidium pullulans Maintain Low Intracellular Cation Concentrations in Hypersaline Environments. Applied and environmental microbiology 71 (11), 6600–6605. 10.1128/AEM.71.11.6600-6605.2005.

Kolar, C.S., Rahel, F.J., 1993. Interaction of a biotic factor (predator presence) and an abiotic factor (low oxygen) as an influence on benthic invertebrate communities. Oecologia 95 (2), 210–219. 10.1007/BF00323492.

Kopylova, E., Noé, L., Touzet, H., 2012. SortMeRNA: Fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics (Oxford, England) 28 (24), 3211–3217. 10.1093/bioinformatics/bts611.

Kozlowski, L.P., 2016. IPC - Isoelectric Point Calculator. Biology direct 11 (1), 55. 10.1186/s13062-016-0159-9.

Kraus, P.R., Heitman, J., 2003. Coping with stress: Calmodulin and calcineurin in model and pathogenic fungi. Biochemical and Biophysical Research Communications 311 (4), 1151–1157. 10.1016/S0006-291X(03)01528-6.

Krogh, A., Larsson, B., Heijne, G. von, Sonnhammer, E.L., 2001. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. Journal of Molecular Biology 305 (3), 567–580. 10.1006/jmbi.2000.4315.

Kushner, D.J., 1978. Life in high salt and solute concentrations: Halophilic bacteria. Microbial life in extreme environments.

Lages, F., Silva-Graça, M., Lucas, C., 1999. Active glycerol uptake is a mechanism underlying halotolerance in yeasts: A study of 42 species. Microbiology (Reading, England) 145 (Pt 9), 2577–2585. 10.1099/00221287-145-9-2577.

Lai, E.C., 2004. Notch signaling: Control of cell communication and cell fate. Development (Cambridge, England) 131 (5), 965–973. 10.1242/dev.01074.

Lamark, T., Kaasen, I., Eshoo, M.W., Falkenberg, P., McDougall, J., Strøm, A.R., 1991. DNA sequence and analysis of the bet genes encoding the osmoregulatory choline?: Glycine betaine pathway of Escherichia coli. Mol Microbiol 5 (5), 1049–1064. 10.1111/j.1365-2958.1991.tb01877.x.

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

Lanyi, J.K., 1974. Salt-dependent properties of proteins from extremely halophilic bacteria. Bacteriological Reviews 38 (3), 272.

Law, C.W., Chen, Y., Shi, W., Smyth, G.K., 2014. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome biology 15 (2), R29. 10.1186/gb-2014-15-2-r29.

Lentendu, G., Buosi, P.R.B., Cabral, A.F., Trevizan Segóvia, B., Ramos Meira, B., Lansac-Tôha, F.M., Velho, L.F.M., Ritter, C.D., Dunthorn, M., 2018. Protist Biodiversity and Biogeography in Lakes From Four Brazilian River-Floodplain Systems. The Journal of eukaryotic microbiology. 10.1111/jeu.12703.

Li, B., Dewey, C.N., 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC bioinformatics 12, 323. 10.1186/1471-2105-12-323.

Li, L., Kim, B.-G., Cheong, Y.H., Pandey, G.K., Luan, S., 2006. A Ca(2)+ signaling pathway regulates a K(+) channel for low-K response in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America 103 (33), 12625–12630. 10.1073/pnas.0605129103.

Li, W., Godzik, A., 2006. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics (Oxford, England) 22 (13), 1658–1659. 10.1093/bioinformatics/btl158.

Liu, Z., Hu, S., Caron, D., 2018. EukZoo, an aquatic protistan protein database for meta-omics studies.

Logares, R., Bråte, J., Bertilsson, S., Clasen, J.L., Shalchian-Tabrizi, K., Rengefors, K., 2009. Infrequent marine-freshwater transitions in the microbial world. Trends in microbiology 17 (9), 414–422. 10.1016/j.tim.2009.05.010.

Lomolino, M.V., Brown, J.H., Sax, D.F., 2010. Island biogeography theory. The theory of island biogeography revisited 13.

Lomolino, M.V., Riddle, B.R., Whittaker, R.J., 2017. Biogeography: Biological diversity across space and time. Oxford University Press Inc, New York, 759 Seiten.

Loomis, W.F., Kuspa, A., Shaulsky, G., 1998. Two-component signal transduction systems in eukaryotic microorganisms. Current opinion in microbiology 1 (6), 643–648. 10.1016/S1369-5274(98)80109-4.

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology 15 (12), 550. 10.1186/s13059-014-0550-8.

Lozupone, C.A., Knight, R., 2007. Global patterns in bacterial diversity. Proceedings of the National Academy of Sciences of the United States of America 104 (27), 11436–11440. 10.1073/pnas.0611525104.

Madern, D., Ebel, C., Zaccai, G., 2000. Halophilic adaptation of enzymes. Extremophiles 4 (2), 91–98. 10.1007/s007920050142.

Maeda, T., Wurgler-Murphy, S.M., Saito, H., 1994. A two-component system that regulates an osmosensing MAP kinase cascade in yeast. Nature 369 (6477), 242–245. 10.1038/369242a0.

Mahé, F., Rognes, T., Quince, C., Vargas, C. de, Dunthorn, M., 2015. Swarm v2: Highly-scalable and high-resolution amplicon clustering. PeerJ 3, e1420. 10.7717/peerj.1420.

Martinac, B., Kloda, A., 2003. Evolutionary origins of mechanosensitive ion channels. Progress in Biophysics and Molecular Biology 82 (1-3), 11–24. 10.1016/S0079-6107(03)00002-6.

Martindale, D.W., 1989. Codon Usage in Tetrahymena and Other Ciliates. The Journal of Protozoology 36 (1), 29–34. 10.1111/j.1550-7408.1989.tb02679.x.

Martínez-Álvarez, R.M., Morales, A.E., Sanz, A., 2005. Antioxidant Defenses in Fish: Biotic and Abiotic Factors. Rev Fish Biol Fisheries 15 (1-2), 75–88. 10.1007/s11160-005-7846-4.

Masutani, H., Yodoi, J., 2001. Overview, in: Sies, H., Packer, L., Colowick, S.P., Kaplan, N.O., Abelson, J.N., Simon, M.I. (Eds.), Methods in enzymology, vol. 347. Academic Press, San Diego, Calif., London, pp. 279–286.

McBride, H.M., Neuspiel, M., Wasiak, S., 2006. Mitochondria: More than just a powerhouse. Current biology : CB 16 (14), R551-60. 10.1016/j.cub.2006.06.054.

Mead, A., Carlton, J.T., Griffiths, C.L., Rius, M., 2011. Introduced and cryptogenic marine and estuarine species of South Africa. Journal of Natural History 45 (39-40), 2463–2524. 10.1080/00222933.2011.595836.

Médicis, E. de, Paquette, J., Gauthier, J.-J., Shapcott, D., 1986. Magnesium and manganese content of halophilic bacteria. Applied and environmental microbiology 52 (3), 567–573.

Medlin, L., Elwood, H.J., Stickel, S., Sogin, M.L., 1988. The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. Gene 71 (2), 491–499.

Meier, H.E.M., Kjellström, E., Graham, L.P., 2006. Estimating uncertainties of projected Baltic Sea salinity in the late 21st century. Geophys. Res. Lett. 33 (15), 327. 10.1029/2006GL026488.

Micklin, P., 2007. The Aral Sea Disaster. Annu. Rev. Earth Planet. Sci. 35 (1), 47–72. 10.1146/annurev.earth.35.031306.140120.

Muallem, S., Schoeffield, M., Pandol, S., Sachs, G., 1985. Inositol trisphosphate modification of ion transport in rough endoplasmic reticulum. Proceedings of the National Academy of Sciences of the United States of America 82 (13), 4433–4437. 10.1073/pnas.82.13.4433.

Niu, B., Zhu, Z., Fu, L., Wu, S., Li, W., 2011. FR-HIT, a very fast program to recruit metagenomic reads to homologous reference genomes. Bioinformatics (Oxford, England) 27 (12), 1704–1705. 10.1093/bioinformatics/btr252.

Nübel, U., Garcia-Pichel, F., Clavero, E., Muyzer, G., 2000. Matching molecular diversity and ecophysiology of benthic cyanobacteria and diatoms in communities along a salinity gradient. Environ Microbiol 2 (2), 217–226. 10.1046/j.1462-2920.2000.00094.x.

Nyyssölä, A., Kerovuo, J., Kaukinen, P., Weymarn, N. von, Reinikainen, T., 2000. Extreme halophiles synthesize betaine from glycine by methylation. The Journal of biological chemistry 275 (29), 22196–22201. 10.1074/jbc.M910111199.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Wagner, H., 2015. Vegan: Community ecology package. R package vegan, vers. 2.2-1.

O'Leary, N.A., Wright, M.W., Brister, J.R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., Astashyn, A., Badretdin, A., Bao, Y., Blinkova, O., Brover, V., Chetvernin, V., Choi, J., Cox, E., Ermolaeva, O., Farrell, C.M., Goldfarb, T., Gupta, T., Haft, D., Hatcher, E., Hlavina, W., Joardar, V.S., Kodali, V.K., Li, W., Maglott, D., Masterson, P., McGarvey, K.M., Murphy, M.R., O'Neill, K., Pujar, S., Rangwala, S.H., Rausch, D., Riddick, L.D., Schoch, C., Shkeda, A., Storz, S.S., Sun, H., Thibaud-Nissen, F., Tolstoy, I., Tully, R.E., Vatsan, A.R., Wallin, C., Webb, D., Wu, W., Landrum, M.J., Kimchi, A., Tatusova, T., DiCuccio, M., Kitts, P., Murphy, T.D., Pruitt, K.D., 2016. Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. Nucleic acids research 44 (D1), D733-45. 10.1093/nar/gkv1189.

Oliver, S., 2000. Guilt-by-association goes global. Nature 403 (6770), 601–603. 10.1038/35001165.

Oren, A., 1999a. Bioenergetic aspects of halophilism. Microbiology and molecular biology reviews : MMBR 63 (2), 334–348.

Oren, A., 1999b. Microbiological studies in the Dead Sea: Future challenges toward the understanding of life at the limit of salt concentrations. Hydrobiologia 405 (0), 1–9. 10.1023/A:1003879932328.

Oren, A., 2001. The bioenergetic basis for the decrease in metabolic diversity at increasing salt concentrations: Implications for the functioning of salt lake ecosystems. Hydrobiologia 466 (1), 61–72. 10.1023/A:1014557116838.

Oren, A., 2002a. Diversity of halophilic microorganisms: Environments, phylogeny, physiology, and applications. Journal of industrial microbiology & biotechnology 28 (1), 56–63. 10.1038/sj/jim/7000176.

Oren, A., 2002b. Molecular ecology of extremely halophilic Archaea and Bacteria. FEMS microbiology ecology 39 (1), 1–7. 10.1111/j.1574-6941.2002.tb00900.x.

Oren, A., 2005. A hundred years of Dunaliella research: 1905-2005. Saline systems 1, 2. 10.1186/1746-1448-1-2.

Oren, A., 2006. Life at High Salt Concentrations, in: Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K.-H., Stackebrandt, E. (Eds.), The Prokaryotes. Volume 2: Ecophysiology and Biochemistry. Springer-Verlag, New York, NY, pp. 263–282.

Oren, A., 2008. Microbial life at high salt concentrations: phylogenetic and metabolic diversity. Saline systems 4 (1), 2. 10.1186/1746-1448-4-2.

Oren, A., 2013. Life at high salt concentrations, intracellular KCl concentrations, and acidic proteomes. Front. Microbiol. 4. 10.3389/fmicb.2013.00315.

Oren, A., Heldal, M., Norland, S., Galinski, E., 2002. Intracellular ion and organic solute concentrations of the extremely halophilic bacterium Salinibacter ruber. Extremophiles 6 (6), 491–498. 10.1007/s00792-002-0286-3.

Patterson, D.J., 1999. The Diversity of Eukaryotes. The American Naturalist 154 (S4), S96-S124. 10.1086/303287.

Paul, S., Bag, S.K., Das, S., Harvill, E.T., Dutta, C., 2008. Molecular signature of hypersaline adaptation: Insights from genome and proteome composition of halophilic prokaryotes. Genome biology 9 (4), R70. 10.1186/gb-2008-9-4-r70.

Pavloudi, C., Kristoffersen, J.B., Oulas, A., Troch, M. de, Arvanitidis, C., 2017. Sediment microbial taxonomic and functional diversity in a natural salinity gradient challenge Remane's "species minimum" concept. PeerJ 5, e3687. 10.7717/peerj.3687.

Pedrós-Alió, Calderón-Paz, MacLean, Medina, Marrasé, Gasol, Guixa-Boixereu, 2000. The microbial food web along salinity gradients. FEMS microbiology ecology 32 (2), 143–155. 10.1111/j.1574-6941.2000.tb00708.x.

Pedrós-Alió, C., 2005. Diversity of Microbial Communities: The Case of Solar Salterns, in: Adaptation to Life at High Salt Concentrations in Archaea, Bacteria, and Eukarya. Springer Netherlands, Dordrecht, pp. 71–90.

Perlman, H., U.S. Geological Survey. How much water is there on Earth, from the USGS Water Science School. USGS - U.S. Geological Survey. https://water.usgs.gov/edu/earthhowmuch.html. Accessed 4 February 2019.

Pertola, S., Kuosa, H., Olsonen, R., 2005. Is the invasion of Prorocentrum minimum (Dinophyceae) related to the nitrogen enrichment of the Baltic Sea? Harmful Algae 4 (3), 481–492. 10.1016/j.hal.2004.08.005.

Petersen, T.N., Brunak, S., Heijne, G. von, Nielsen, H., 2011. SignalP 4.0: Discriminating signal peptides from transmembrane regions. Nature Methods 8 (10), 785–786. 10.1038/nmeth.1701.

Petrovicˇ, U., Gunde-Cimerman, N., Plemenitasˇ, A., 2002. Cellular responses to environmental salinity in the halophilic black yeast Hortaea werneckii. Molecular microbiology 45 (3), 665–672.

Pinto, A.J., Raskin, L., 2012. PCR biases distort bacterial and archaeal community structure in pyrosequencing datasets. PloS one 7 (8), e43093. 10.1371/journal.pone.0043093.

Piscart, C., Moreteau, J.-C., Beisel, J.-N., 2005. Biodiversity and Structure of Macroinvertebrate Communities Along a Small Permanent Salinity Gradient (Meurthe River, France). Hydrobiologia 551 (1), 227–236. 10.1007/s10750-005-4463-0.

Plemenitaš, A., Lenassi, M., Konte, T., Kejžar, A., Zajc, J., Gostinčar, C., Gunde-Cimerman, N., 2014. Adaptation to high salt concentrations in halotolerant/halophilic fungi: A molecular perspective. Frontiers in microbiology 5, 199. 10.3389/fmicb.2014.00199.

Prescott, D.M., 1994. The DNA of ciliated protozoa. Microbiol. Mol. Biol. Rev. 58 (2), 233–267.

Prescott, D.M., Prescott, J.D., Prescott, R.M., 2002. Coding properties of macronuclear DNA molecules in Sterkiella nova (Oxytricha nova). Protist 153 (1), 71–77. 10.1078/1434-4610-00084.

Prista, C., Almagro, A., Loureiro-Dias, M.C., Ramos, J., 1997. Physiological basis for the high salt tolerance of Debaryomyces hansenii. Applied and environmental microbiology 63 (10), 4005–4009.

Prista, C., Loureiro-Dias, M.C., Montiel, V., García, R., Ramos, J., 2005. Mechanisms underlying the halotolerant way of Debaryomyces hansenii. FEMS yeast research 5 (8), 693–701.

R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Ramos, J., 2005. Introducing Debaryomyces hansenii, a salt loving yeast. Adaptation to Life at High Salt Concentrations in Archaea, Bacteria, and Eukarya 9, 443.

Ramos, J., Ariño, J., Sychrová, H., 2011. Alkali-metal-cation influx and efflux systems in nonconventional yeast species. FEMS Microbiology Letters 317 (1), 1–8. 10.1111/j.1574-6968.2011.02214.x.

Rathinasabapathi, B., Burnet, M., Russell, B.L., Gage, D.A., Liao, P.-C., Nye, G.J., Scott, P., Golbeck, J.H., Hanson, A.D., 1997. Choline monooxygenase, an unusual iron-sulfur enzyme catalyzing the first step of glycine betaine synthesis in plants: Prosthetic group characterization and cDNA cloning. Proceedings of the National Academy of Sciences of the United States of America 94 (7), 3454–3458. 10.1073/pnas.94.7.3454.

Reed, R.H., Chudek, J.A., Foster, R., Stewart, W.D.P., 1984. Osmotic adjustment in cyanobacteria from hypersaline environments. Arch. Microbiol. 138 (4), 333–337. 10.1007/BF00410900.

Rhodes, M.E., Fitz-Gibbon, S.T., Oren, A., House, C.H., 2010. Amino acid signatures of salinity on an environmental scale with a focus on the Dead Sea. Environmental microbiology 12 (9), 2613–2623. 10.1111/j.1462-2920.2010.02232.x.

Rivera, S.F., Vasselon, V., Ballorain, K., Carpentier, A., Wetzel, C.E., Ector, L., Bouchez, A., Rimet, F., 2018. DNA metabarcoding and microscopic analyses of sea turtles biofilms: Complementary to understand turtle behavior. PloS one 13 (4), e0195770. 10.1371/journal.pone.0195770.

Roberts, M.F., 2004. Osmoadaptation and osmoregulation in archaea: Update 2004. Frontiers in Bioscience 9, 1999–2019.

Roberts, M.F., 2005. Organic compatible solutes of halotolerant and halophilic microorganisms. Saline Syst 1 (1), 5. 10.1186/1746-1448-1-5.

Robinson, M.D., Oshlack, A., 2010. A scaling normalization method for differential expression analysis of RNA-seq data. Genome biology 11 (3), R25. 10.1186/gb-2010-11-3-r25.

Rodríguez-Navarro, A., Benito, B., 2010. Sodium or potassium efflux ATPase a fungal, bryophyte, and protozoal ATPase. Biochimica et Biophysica Acta 1798 (10), 1841–1853. 10.1016/j.bbamem.2010.07.009.

Roeßler, M., Müller, V., 2001. Osmoadaptation in bacteria and archaea: Common principles and differences. Environ Microbiol 3 (12), 743–754. 10.1046/j.1462-2920.2001.00252.x.

Rognes, T., Flouri, T., Nichols, B., Quince, C., Mahé, F., 2016. VSEARCH: A versatile open source tool for metagenomics. PeerJ 4, e2584. 10.7717/peerj.2584.

Rosenberg, R., Loo, L.-O., Möller, P., 1992. Hypoxia, salinity and temperature as structuring factors for marine benthic communities in a eutrophic area. Netherlands Journal of Sea Research 30, 121–129. 10.1016/0077-7579(92)90051-F.

Ruinen, J., 1938. Notizen über Ciliaten aus konzentrierten Salgewässern. Zool. Meded. Leiden (20), 243–256.

Saier, M.H., Reddy, V.S., Tsu, B.V., Ahmed, M.S., Li, C., Moreno-Hagelsieb, G., 2016. The Transporter Classification Database (TCDB): Recent advances. Nucleic acids research 44 (D1), D372-9. 10.1093/nar/gkv1103.

Saier, M.H., Tran, C.V., Barabote, R.D., 2006. TCDB: The Transporter Classification Database for membrane transport protein analyses and information. Nucleic acids research 34 (Database issue), D181-6. 10.1093/nar/gkj001.

Santos, H., da Costa, M.S., 2002. Compatible solutes of organisms that live in hot saline environments. Environ Microbiol 4 (9), 501–509. 10.1046/j.1462-2920.2002.00335.x.

Saum, S.H., Müller, V., 2008. Regulation of osmoadaptation in the moderate halophile Halobacillus halophilus: Chloride, glutamate and switching osmolyte strategies. Saline systems 4, 4. 10.1186/1746-1448-4-4.

Singer, E., Bushnell, B., Coleman-Derr, D., Bowman, B., Bowers, R.M., Levy, A., Gies, E.A., Cheng, J.-F., Copeland, A., Klenk, H.-P., Hallam, S.J., Hugenholtz, P., Tringe, S.G., Woyke, T., 2016. High-resolution phylogenetic microbial community profiling. The ISME journal 10 (8), 2020–2032. 10.1038/ismej.2015.249.

Skarlato, S., Filatova, N., Knyazev, N., Berdieva, M., Telesh, I., 2018. Salinity stress response of the invasive dinoflagellate Prorocentrum minimum. Estuarine, Coastal and Shelf Science 211, 199–207. 10.1016/j.ecss.2017.07.007.

Smith-Unna, R., Boursnell, C., Patro, R., Hibberd, J.M., Kelly, S., 2016. TransRate: reference-free quality assessment of de novo transcriptome assemblies. Genome research 26 (8), 1134–1144. 10.1101/gr.196469.115.

Sommer, U., Adrian, R., Senerpont Domis, L. de, Elser, J.J., Gaedke, U., Ibelings, B., Jeppesen, E., Lürling, M., Molinero, J.C., Mooij, W.M., van Donk, E., Winder, M., 2012. Beyond the Plankton Ecology Group (PEG) Model: Mechanisms Driving Plankton Succession. Annu. Rev. Ecol. Evol. Syst. 43 (1), 429–448. 10.1146/annurev-ecolsys-110411-160251.

Sommer, U., Gliwicz, Z.M., Lampert, W., Duncan, A., 1986. The PEG-model of seasonal succession of planktonic events in fresh waters. Arch. Hydrobiol 106 (4), 433–471.

Sonnhammer, E.L., Heijne, G. von, Krogh, A., 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. Proceedings. International Conference on Intelligent Systems for Molecular Biology 6, 175–182.

Soppa, J., 2006. From genomes to function: haloarchaea as model organisms. Microbiology (Reading, England) 152 (Pt 3), 585–590. 10.1099/mic.0.28504-0.

Sorensen, K.B., Canfield, D.E., Oren, A., 2004. Salinity Responses of Benthic Microbial Communities in a Solar Saltern (Eilat, Israel). Applied and environmental microbiology 70 (3), 1608–1616. 10.1128/AEM.70.3.1608-1616.2004.

Stephens, D.W., 1990. Changes in lake levels, salinity and the biological community of Great Salt Lake (Utah, USA), 1847–1987, in: Comín, F.A., Northcote, T.G. (Eds.), Saline Lakes. Proceedings of the Fourth International Symposium on Athalassic (inland) Saline Lakes, held at Banyoles, Spain, May 1988. Springer Netherlands, Dordrecht, pp. 139–146.

Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M.D.M., Breiner, H.-W., Richards, T.A., 2010. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. Molecular ecology 19 Suppl 1, 21–31. 10.1111/j.1365-294X.2009.04480.x.

Stoeck, T., Breiner, H.-W., Filker, S., Ostermaier, V., Kammerlander, B., Sonntag, B., 2014. A morphogenetic survey on ciliate plankton from a mountain lake pinpoints the necessity of lineage-specific barcode markers in microbial ecology. Environmental microbiology 16 (2), 430–444. 10.1111/1462-2920.12194.

Suescún-Bolívar, L.P., Thomé, P.E., 2015. Osmosensing and osmoregulation in unicellular eukaryotes. World journal of microbiology & biotechnology 31 (3), 435–443. 10.1007/s11274-015-1811-8.

Suzuki, S., Kakuta, M., Ishida, T., Akiyama, Y., 2014. GHOSTX: An improved sequence homology search algorithm using a query suffix array and a database suffix array. PloS one 9 (8), e103833. 10.1371/journal.pone.0103833.

Taylor, F.J.R., Hoppenrath, M., Saldarriaga, J.F., 2008. Dinoflagellate diversity and distribution. Biodivers Conserv 17 (2), 407–418. 10.1007/s10531-007-9258-3.

Telesh, I., Postel, L., Heerkloss, R., Mironova, E., Skarlato, S., 2008. Zooplankton of the Open Baltic Sea: Atlas.

Telesh, I.V., Schubert, H., Skarlato, S.O., 2011. Revisiting Remane's concept: Evidence for high plankton diversity and a protistan species maximum in the horohalinicum of the Baltic Sea. Mar. Ecol. Prog. Ser. 421, 1–11. 10.3354/meps08928.

Telesh, I.V., Schubert, H., Skarlato, S.O., 2013. Life in the salinity gradient: Discovering mechanisms behind a new biodiversity pattern. Estuarine, Coastal and Shelf Science 135, 317–327. 10.1016/j.ecss.2013.10.013.

Telesh, I.V., Schubert, H., Skarlato, S.O., 2016. Ecological niche partitioning of the invasive dinoflagellate Prorocentrum minimum and its native congeners in the Baltic Sea. Harmful Algae 59, 100–111. 10.1016/j.hal.2016.09.006.

Tittensor, D.P., Mora, C., Jetz, W., Lotze, H.K., Ricard, D., Berghe, E.V., Worm, B., 2010. Global patterns and predictors of marine biodiversity across taxa. Nature 466 (7310), 1098–1101. 10.1038/nature09329.

Toone, W.M., Morgan, B.A., Jones, N., 2001. Redox control of AP-1-like factors in yeast and beyond. Oncogene 20 (19), 2336–2346. 10.1038/sj.onc.1204384.

Triadó-Margarit, X., Casamayor, E.O., 2013. High genetic diversity and novelty in planktonic protists inhabiting inland and coastal high salinity water bodies. FEMS microbiology ecology 85 (1), 27–36. 10.1111/1574-6941.12095.

Tuomisto, H., 2010. A diversity of beta diversities: Straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. Ecography 33 (1), 2–22. 10.1111/j.1600-0587.2009.05880.x.

UniProt Consortium, T., 2017. UniProt: The universal protein knowledgebase. Nucleic acids research 45 (D1), D158-D169. 10.1093/nar/gkw1099.

United Nations. 2010–2020: UN Decade for Deserts and the Fight against Desertification. https://www.un.org/en/events/desertification_decade/whynow.shtml. Accessed 24 April 2019.

United Nations, 2017. Factsheet: People and Oceans. The Ocean Conference. United Nations. https://www.un.org/sustainabledevelopment/wp-content/uploads/2017/05/Ocean-fact-sheet-package.pdf. Accessed 24 April 2019.

Vargas, C. de, Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Berney, C., Le Bescot, N., Probert, I., 2015. Eukaryotic plankton diversity in the sunlit ocean. Science (New York, N.Y.) 348 (6237), 1261605.

Veness, C., 2002-2017. Calculate distance and bearing between two Latitude/Longitude points using haversine formula in JavaScript. https://www.movable-type.co.uk/scripts/latlong.html. Accessed 10 January 2019.

Verschuren, D., Tibby, J., Sabbe, K., Roberts, N., 2000. Effects of Depth, Salinity, and Substrate on the Invertebrate Community of a Fluctuating Tropical Lake. Ecology 81 (1), 164. 10.2307/177142.

Visco, J.A., Apothéloz-Perret-Gentil, L., Cordonier, A., Esling, P., Pillet, L., Pawlowski, J., 2015. Environmental Monitoring: Inferring the Diatom Index from Next-Generation Sequencing Data. Environmental science & technology 49 (13), 7597–7605. 10.1021/es506158m.

Wang, J., Yang, D., Zhang, Y., Shen, J., van der Gast, C., Hahn, M.W., Wu, Q., 2011. Do patterns of bacterial diversity along salinity gradients differ from those observed for macroorganisms? PloS one 6 (11), e27597. 10.1371/journal.pone.0027597.

Wasmund, N., Dutz, J., Pollehne, F., Siegel, H., Zettler, M.L., 2017. Biological assessment of the Baltic Sea 2016 (105). 10.12754/msr-2017-0105.

Wasmund, N., Nausch, G., Feistel, R., 2013. Silicate consumption: An indicator for long-term trends in spring diatom development in the Baltic Sea. Journal of Plankton Research 35 (2), 393–406. 10.1093/plankt/fbs101.

Waterkeyn, A., Grillas, P., Vanschoenwinkel, B., Brendonck, L., 2008. Invertebrate community patterns in Mediterranean temporary wetlands along hydroperiod and salinity gradients. Freshwater Biology 53 (9), 1808–1822. 10.1111/j.1365-2427.2008.02005.x.

Weinisch, L., Kirchner, I., Grimm, M., Kühner, S., Pierik, A.J., Rosselló-Móra, R., Filker, S., 2018a. Glycine Betaine and Ectoine Are the Major Compatible Solutes Used by Four Different Halophilic Heterotrophic Ciliates. Microbial ecology. 10.1007/s00248-018-1230-0.

Weinisch, L., Kühner, S., Roth, R., Grimm, M., Roth, T., Netz, D.J.A., Pierik, A.J., Filker, S., 2018b. Identification of osmoadaptive strategies in the halophile, heterotrophic ciliate Schmidingerothrix salinarum. PLoS biology 16 (1), e2003892. 10.1371/journal.pbio.2003892.

Weiss, M., Pick, U., 1991. Uptake of the fluorescent indicator atebrin into acidic vacuoles in the halotolerant alga Dunaliella salina. Planta 185 (4), 494–501. 10.1007/BF00202958.

Weisse, T., 2008. Distribution and diversity of aquatic protists: An evolutionary and ecological perspective. Biodivers Conserv 17 (2), 243–259. 10.1007/s10531-007-9249-4.

Wells, H.W., 1961. The Fauna of Oyster Beds, with Special Reference to the Salinity Factor. Ecological Monographs 31 (3), 239–266. 10.2307/1948554.

Welsh, D.T., 2000. Ecological significance of compatible solute accumulation by micro-organisms: from single cells to global climate. FEMS Microbiology Reviews 24 (3), 263–290. 10.1111/j.1574-6976.2000.tb00542.x.

Wickham, H., 2016. Ggplot2: Elegant Graphics for Data Analysis, 2nd ed. Springer International Publishing; Imprint; Springer, Cham, 1 online resource (XVI, 260.

Wingfield, J.C., Boonstra, R., 2013. Ecological processes and the ecology of stress: The impacts of abiotic environmental factors. Funct Ecol 27 (1), 37–44. 10.1111/1365-2435.12039.

Witman, J.D., Grange, K.R., 1998. Links between Rain, Salinity, and Predation in a Rocky Subtidal Community. Ecology 79 (7), 2429. 10.2307/176833.

Wu, Q.L., Zwart, G., Schauer, M., Kamst-van Agterveld, M.P., Hahn, M.W., 2006. Bacterioplankton community composition along a salinity gradient of sixteen high-mountain lakes located on the Tibetan Plateau, China. Applied and environmental microbiology 72 (8), 5478–5485. 10.1128/AEM.00767-06.

Wuichet, K., Cantwell, B.J., Zhulin, I.B., 2010. Evolution and phyletic distribution of two-component signal transduction systems. Current opinion in microbiology 13 (2), 219–225. 10.1016/j.mib.2009.12.011.

Wuyts, J., Perrière, G., van de Peer, Y., 2004. The European ribosomal RNA database. Nucleic acids research 32 (Database issue), D101-3. 10.1093/nar/gkh065.

Yancey, P., Clark, M., Hand, S., Bowlus, R., Somero, G., 1982. Living with water stress: Evolution of osmolyte systems. Science 217 (4566), 1214–1222. 10.1126/science.7112124.

Zajc, J., Liu, Y., Dai, W., Yang, Z., Hu, J., Gostinčar, C., Gunde-Cimerman, N., 2013. Genome and transcriptome sequencing of the halophilic fungus Wallemia ichthyophaga: Haloadaptations present and absent. BMC genomics 14, 617. 10.1186/1471-2164-14-617.

Zhang, J., Tan, W., Yang, X.-H., Zhang, H.-X., 2008. Plastid-expressed choline monooxygenase gene improves salt and drought tolerance through accumulation of glycine betaine in tobacco. Plant cell reports 27 (6), 1113–1124. 10.1007/s00299-008-0549-2.

Zhao, H., 2005. Effect of ions and other compatible solutes on enzyme activity, and its implication for biocatalysis using ionic liquids. Journal of Molecular Catalysis B: Enzymatic 37 (1-6), 16–25. 10.1016/j.molcatb.2005.08.007.

Zhu, F., Massana, R., Not, F., Marie, D., Vaulot, D., 2005. Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. FEMS microbiology ecology 52 (1), 79–92. 10.1016/j.femsec.2004.10.006.

VIII. Appendix

The appendix is divided in two different parts: One attached part to this document and another digital part placed on a contributed CD-Rom.

The additional digital part on the CD-Rom includes scripts for 'R' and 'bash' as well as all analysed data for shown tables and figures. The CD-Rom contains namely:

- Dissertation_Steffen_Kühner (digital version)
- copyright and licence agreements with John Wiley and sons
    - John Wiley and Sons Licence Agreement Filker et al. 2017
    - John Wiley and Sons Licence Agreement Filker et al. 2019
    - Screenshots_Copyright Clearance Center_Wiley_Filker et al. 2017
    - Screenshots_Copyright Clearance Center_Wiley_Filker et al. 2019
    - Thank you for your order with RightsLink _ John Wiley and Sons
    - Thank you for your order with RightsLink _ John Wiley and Sons_2

- in folder 'Subject_I_metabarcoding':
    - environmental_parameter_Baltic_Sea.xlsx
    - OTU_table_surface.xlsx
    - Rstudio_commands_statistics.R

    Additional material and supplemental data was published before under doi 10.1111/1462-2920.14502 and can be provided upon request.

- in folder 'Subject_II_meta-transcriptome':
    - Abundance_matrix_TPM
    - Annotationtable_complete_with_KEGG.xlsx
    - Commands_annotations+taxonomic_assignment.docx
    - Commands_deeploc.docx
    - Commands_IPC.docx
    - Commands_ORF_prediction.docx
    - Commands_Qiime.docx
    - Commands_SortMeRNA.docx
    - Commands_transcript_quantification.docx
    - Commands_Trimmomatic.docx

- Commands_Trinity+Transrate+CD-HIT-EST.docx
- Commands_vsearch_clustering.docx
- Input_data_annotated_OGs_categorization.xlsx
- Input_data_Fig.15_amount_of_annotations.xlsx
- Input_data_Fig.16_taxonomic_assignment.xlsx
- Input_data_Fig.17_jackknife_named_nodes_TPM_all.tre
- Input_data_Fig.18_KEGG_normalized_accross_all.xlsx
- Input_data_Fig.19_jackknife_named_nodes_sensing.tre
- Input_data_Fig.20_sensing_funs_normalisiert_gesamt.xlsx
- Input_data_Fig.21_jackknife_named_nodes_transport_cs.tre
- Input_data_Fig.22_TPM_transport_cs_funs_normalisiert_gesamt.xlsx
- Input_data_Fig.23_jackknife_named_nodes_compsol.tre
- Input_data_Fig.24_TPM_compsol_funs_normalisiert_gesamt.xlsx
- Input_data_Fig.25_jackknife_named_nodes_transport_ion.tre
- Input_data_Fig.26_transport_ion_funs_normalisiert_gesamt.xlsx
- Input_data_Fig.27_jackknife_named_nodes_energy.tre
- Input_data_Fig.28_TPM_energy_funs_normalisiert_gesamt.xlsx
- Input_for_taxonomic_assignment.xlsx
- Raw_data_Fig.18_KEGG_annotation.xlsx
- Rstudio_commands_data_rearrangements.R
- Rstudio_commands_heatmaps.R
- Rstudio_commands_IPC+deeploc.R
- Rstudio_commands_taxonomic_assignment.R
- Rstudio_commands_TPM.R

- in folder 'Subject_III_transcriptome_S.salinarum':
  - Annotationtable_15093_diffexpr_ORFs.xlsx
  - Annotationtable_168454_ORFs.xlsx
  - Annotationtable_all_clustered_202.480_ORFs_sorted.xlsx
  - Commands_annotations.docx
  - Commands_codon_usage.docx
  - Commands_deeploc.docx
  - Commands_gc_content.docx

- Commands_IPC.docx
- Commands_ORF_prediction.docx
- Commands_PtR_script.docx
- Commands_SortMeRNA.docx
- Commands_transcript_quantification_DEA.docx
- Commands_Transfuse.docx
- Commands_Transrate.docx
- Commands_Trimmomatic.docx
- Commands_Trinity.docx
- Input_data_Fig.30_completeness_KEGG.xlsx
- Input_data_Fig.31_PtR_script.matrix
- Input_data_Fig.32_PtR_script.matrix
- Input_data_Fig.32_regulation_15093_diffexpr_ORFs.xlsx
- Input_data_Fig.33_KEGG.xlsx
- Input_data_Fig.34_deeploc.xlsx
- Input_data_Fig.35_sensing.xlsx
- Input_data_Fig.36+38_transport.xlsx
- Input_data_Fig.37_compatible_solute_synthesis.xlsx
- Input_data_Fig.39_other_processes.xlsx
- Input_data_Fig.40_IPC_Frequencies_species.xlsx
- Input_data_Fig.41_IPC_Frequencies_deeploc.xlsx
- Rstudio_commands_codon_usage.R
- Rstudio_commands_data_rearrangements.R
- Rstudio_commands_deeploc_IPC.R
- Rstudio_commands_heatmaps_Fig.35-39.R
- Rstudio_commands_regulation_matrix.R

All '.docx-files' can be opend with 'Word' and all '.xlsx-files' with 'Excel'. The 'R'-scripts (.R-files) can be opend by 'Tinn-R' or using 'R' itself. The software 'Tinn-R' can be downloaded from 'https://sourceforge.net/projects/tinn-r/' (without warranty).

In the attached appendix below, eight supplemental figures and four supplemental tables can be found.
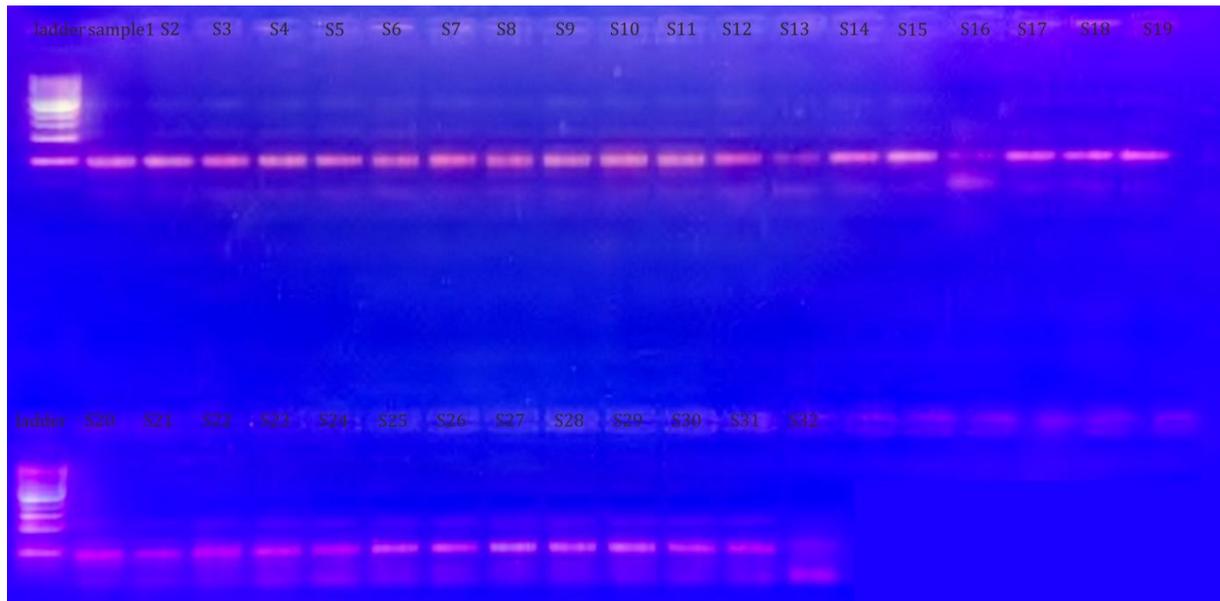
Supplemental Figure 1: Example results of agarose gel electrophoresis of PCR for amplifying the hypervariable V4 region of the SSU rRNA. All samples (S1 – S32) show a clear band at approximately 500 bp.
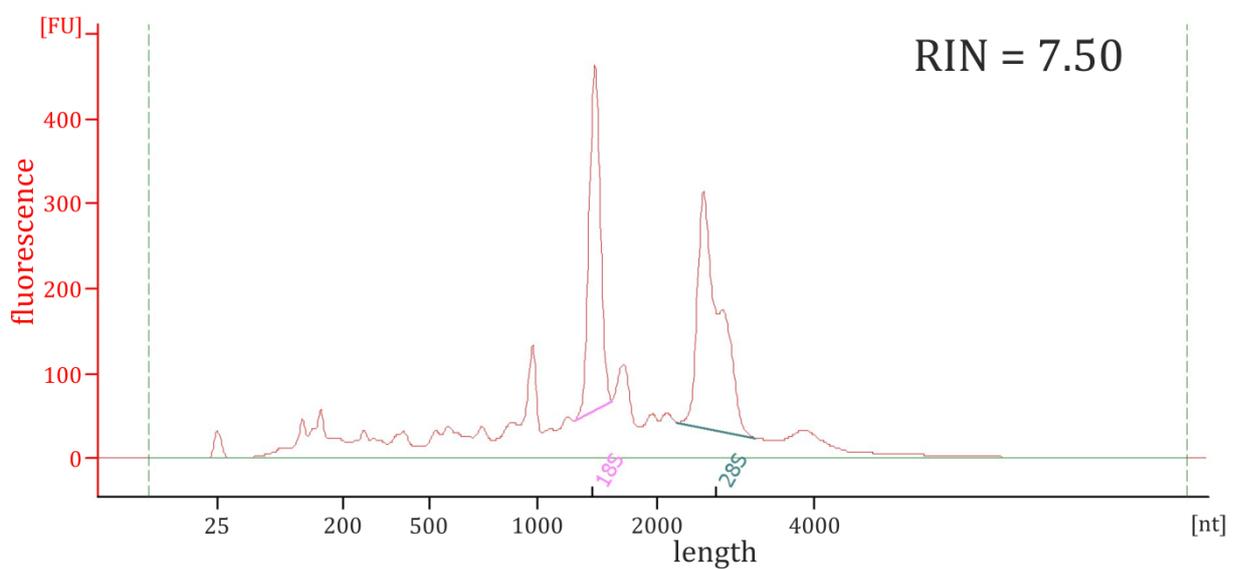
Supplemental Figure 2: Exemplary result of a RNA 6000 PicoAssay on an Agilent 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany). The RIN-number summarized the integrity of the RNA sample and therefore, it is a dimension for the quality of the RNA sample. A RIN-number of 10 means 'intact' and a RIN-number of 0 represents a 'fully degraded' RNA.
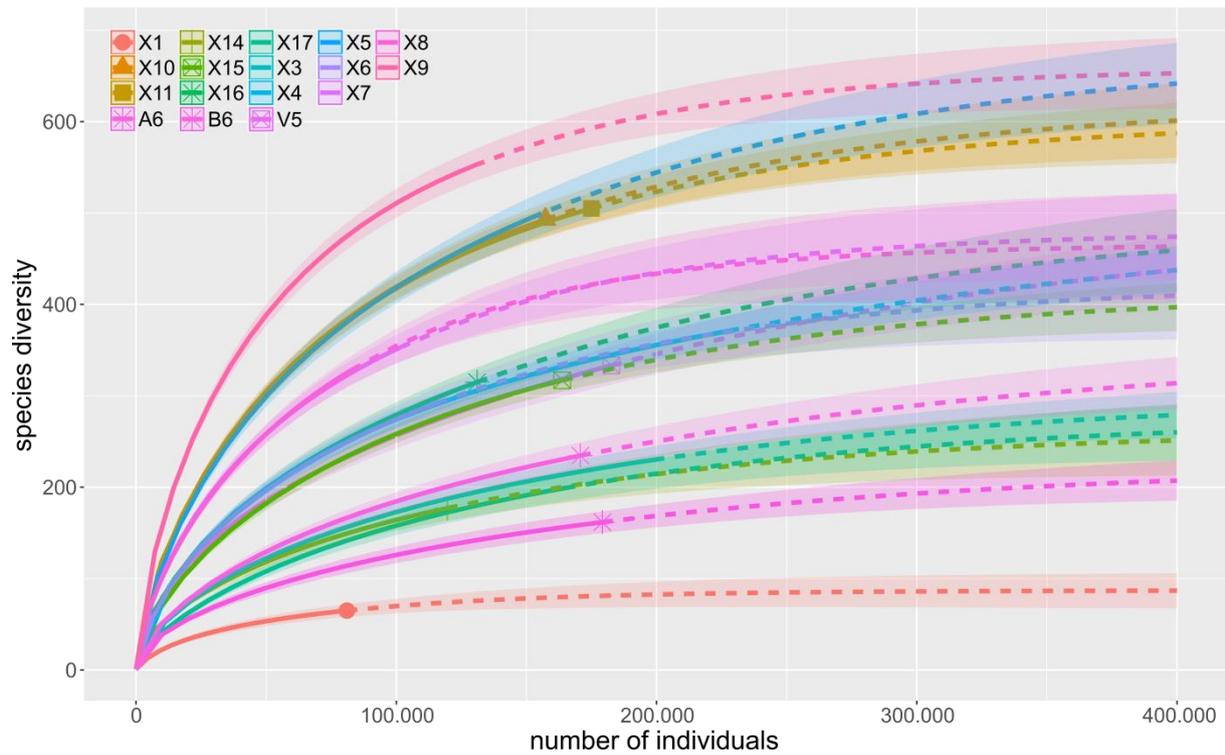
Supplemental Figure 3: Sample-size based rarefaction (solid line) and extrapolation (dotted line) sampling curves for species richness (Hill number q = 0), with 95 % confidence intervals (shaded area) for surface water samples (6 m depth) computed with the R package 'iNext' (Hsieh et al., 2016). The majority of samples approached the asymptote, indicating near-saturated sampling. Measurements were considered for uneven sample sizes/saturation depth in statistical community comparisons (see chapter '1.4 Statistical analyses'). The plot was slightly modified from Filker et al. (2019).
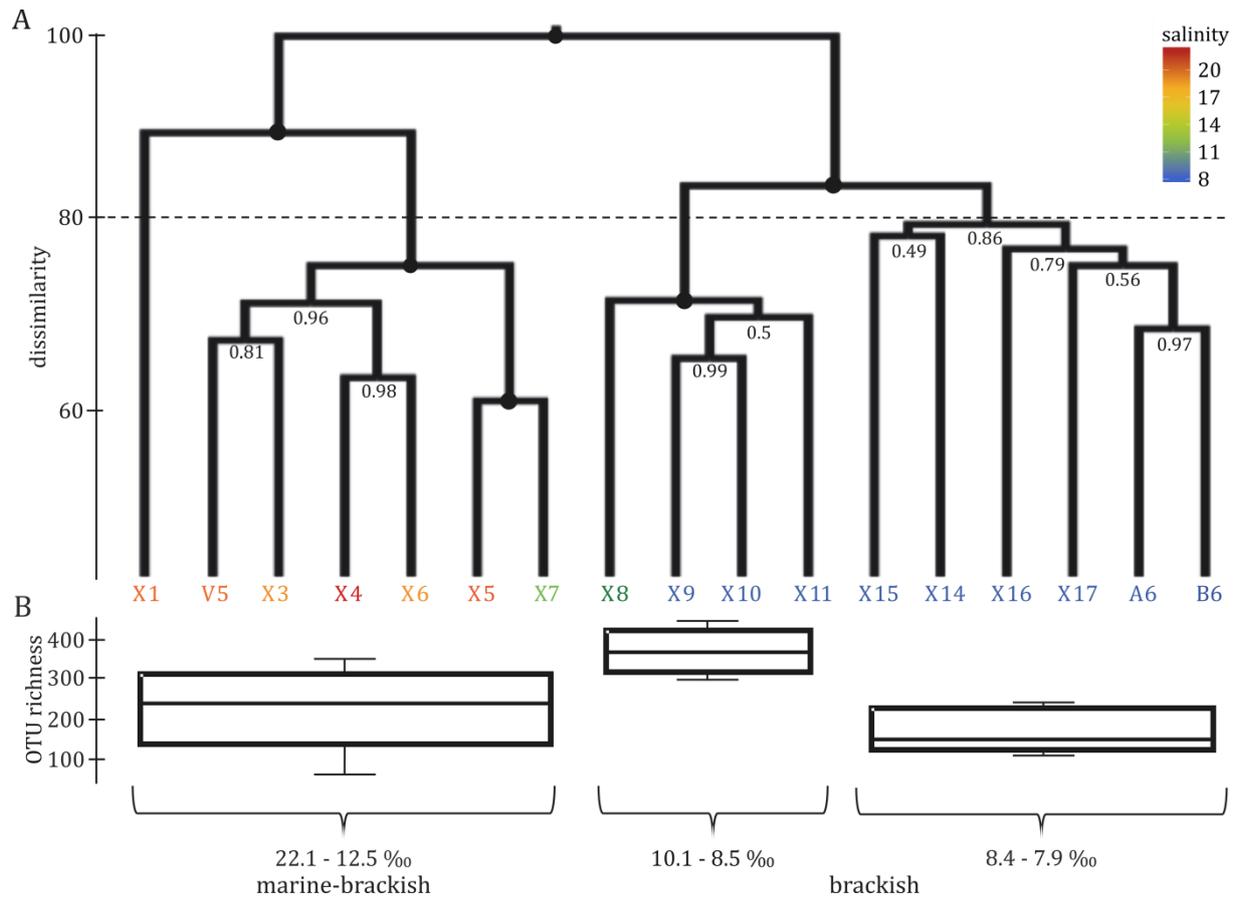
Supplemental Figure 4: A: Hierarchical clustering based on protistan plankton community composition similarities. Samples were collected from surface waters along the west–east transect of the Baltic Sea with a pronounced salinity gradient. Bootstrap values of the dendrogram are shown (• = 1). The protistan plankton communities cluster in three salinities: a marine-brackish cluster (22.1 - 12.5 ‰), and a brackish cluster, which can be further subdivided into one cluster below 8.5 ‰ and one cluster above 8.5 ‰. Differences in communities are more pronounced in the transition from the brackish cluster to the marine-brackish cluster, identifying 10–12 ‰ as the strongest environmental salinity barrier. B: OTU richness within the different salinity clusters. The figure was slightly modified from Filker et al. (2019).
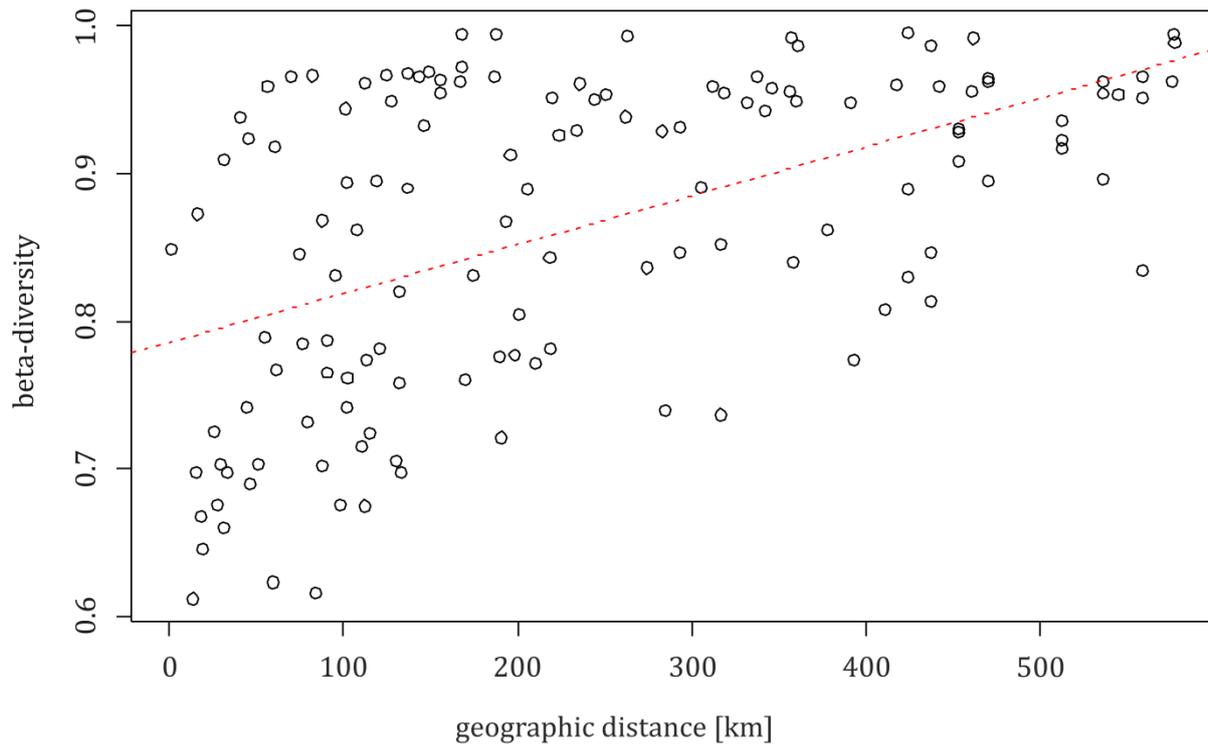
Supplemental Figure 5: Regression plot from Mantel analysis of Bray–Curtis distances as a measure of beta-diversity and geographic distance. While the significance of the correlation between community dissimilarity and salinity differences increased further after controlling for geographic effects ($p < 0.01$, $r = 0.8$), the correlation between geographic distance and community dissimilarity decreased from $r = 0.5$ ($p < 0.01$) to $r = 0.2$ ($p < 0.05$) when controlled for differences in salinity. Thus, geographic distance is no determining factor for structuring of protistan communities in surface water layers along the west–east sampling transect in the Baltic Sea. The plot was slightly modified from Filker et al. (2019).
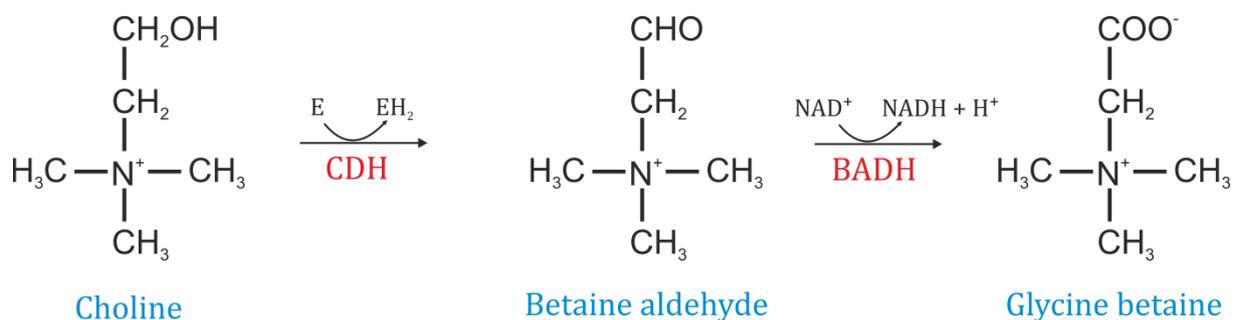
Supplemental Figure 6: Detected pathway of glycine betaine synthesis starting with the precursor choline found in the transcriptome of *Schmidingerothrix salinarum* after induced salt up-shock.
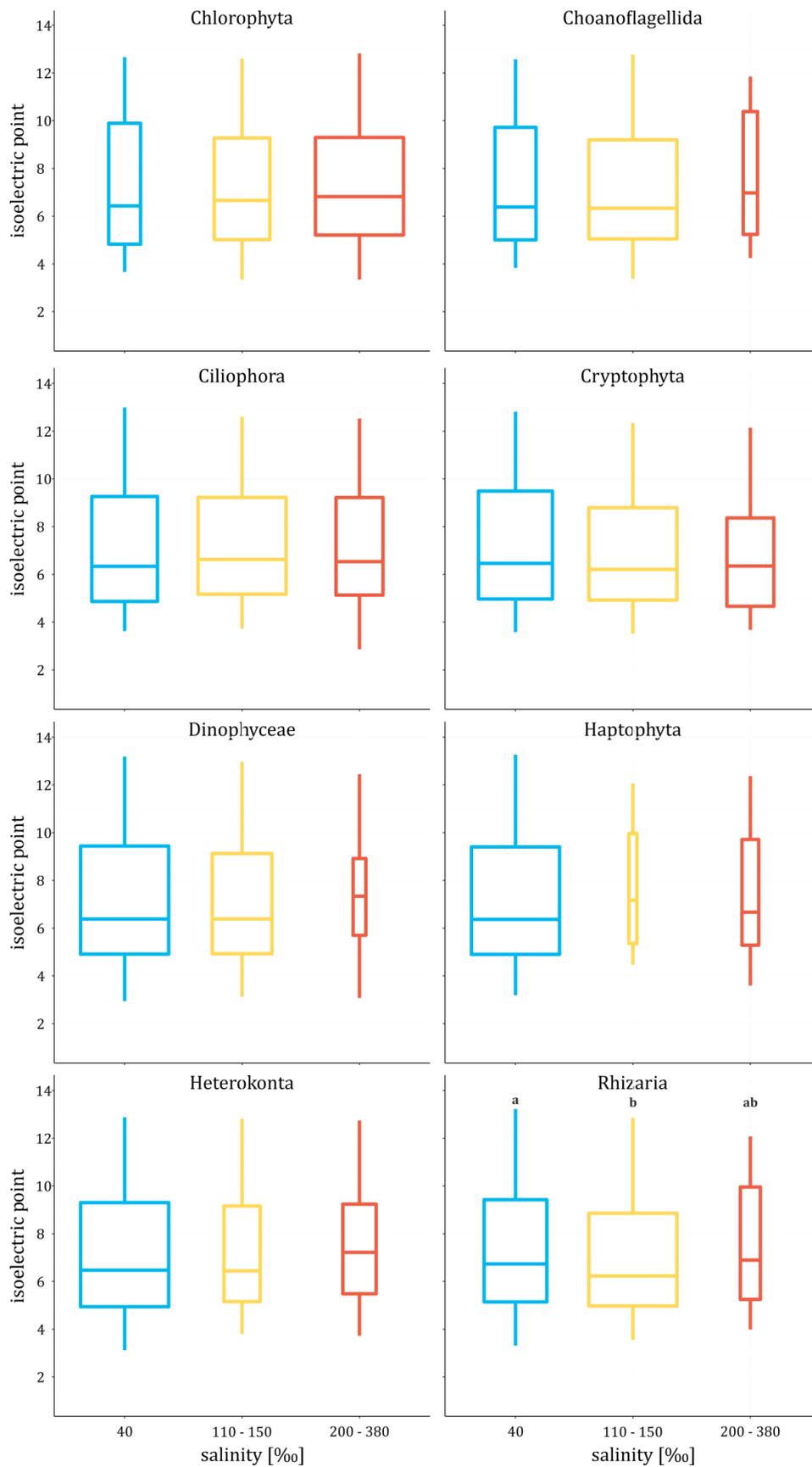
Supplemental Figure 7: Comparing illustration of the IPC results according to different salinities separated after phyla. All boxplots are overlapping each other and the width was related to the amount of proteins used for the calculation. Statistical differences of the mean are symbolised by the usage of different letters (p < 0.05).
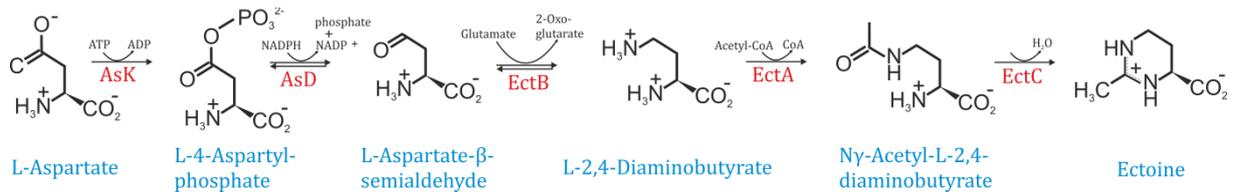
Supplemental Figure 8: Detected pathway of *de novo* synthesis of ectoine starting with the precursor aspartate using the five different enzymes aspartate kinase (AsK), aspartate-semialdehyde dehydrogenase (AsD), diaminobutyrate-2-oxoglutarate aminotransferase (EctB), diaminobutyrate acetytransferase (EctA) and ectoine synthase (EctC), which were found in the transcriptome of *Schmidingerothrix salinarum*.

Supplemental Table 1: Overview of sample collection metadata, including environmental and sequence data information. The table was slightly modified from Filker et al. (2019).

Region	Sub-region	station ID	sampling depth [m]	total depth [m]	coordinates		high-quality target reads	OTUs	salinity [‰]	temperature [°C]	oxygen [ml/l]
Western Baltic Sea	Kiel Bight	X1	6	23	54° 41.25 N	10° 12.08 E	81,007	65	18.38	7.03	8.42
		V5	6	32	54° 41.68 N	10° 20.48 E	182,622	271	18.02	7.01	7.39
		X3	6	26	54° 40.97 N	10° 30.07 E	198,968	230	17.31	7.38	8.46
		X4	6	25	54° 37.00 N	10° 55.07 E	227,612	371	22.05	7.13	8.15
	Mecklenburg Bay	X5	6	27	54° 28.86 N	11° 23.27 E	152,062	495	19.39	7.16	7.74
Baltic Proper	Arkona Basin	X6	6	21	54° 24.01 N	12° 10.21 E	115,564	292	16.75	6.38	7.83
		X7	6	20	54° 36.98 N	12° 17.05 E	101,520	353	12.47	6.67	7.95
		X8	6	18	54° 43.03 N	12° 29.17 E	82,715	328	10.13	5.94	8.20
		X9	6	17	54° 50.02 N	12° 37.39 E	130,024	552	9.23	6.26	8.01
		X10	6	28	54° 54.49 N	12° 47.47 E	157,289	493	8.56	5.50	8.53
		X11	6	44	54° 53.68 N	13° 05.74 E	174,971	505	8.49	5.18	9.36
		A6	6	48	55° 06.57 N	14° 14.48 E	170,579	213	8.02	4.89	8.26
	Bornholm Basin	B6	6	95	55° 17.50 N	15° 44.93 E	179,201	159	7.87	4.78	8.60
		X14	6	89	55° 13.85 N	17° 21.55 E	119,582	177	8.27	4.70	8.72
	Gotland Basin	X15	6	72	54° 47.71 N	15° 44.89 E	163,761	317	8.34	5.05	8.78
		X16	6	114	55° 57.01 N	19° 03.07 E	131,072	315	8.02	4.57	9.14
		X17	6	92	54° 58.18 N	18° 41.95 E	164,744	199	8.39	4.94	9.00

Supplemental Table 2: Slightly modified protocol for DNA/RNA-extraction with AllPrep DNA/RNA Mini Kit (Qiagen, Germany) according to manufacturer's instructions.

1. Cut the filter in 4 pieces and put it in E-matrix-tube (not scrunch but along the margin, that beats can move freely)
2. Add 600 µl RLT Plus buffer
3. Add 6 µl β-mercaptoethanol
4. Centrifuge RNA*later* down (1 min – full speed: 14000 rpm)
5. Solve pelett in 200 µl RLT Plus buffer and 2 µl β-mercaptoethanol and put all together in E-matrix-tube
6. Beat E-matrix-tube for 45 sec at 30 Hz
7. Centrifuge 3 min with full speed
8. Pipet supernatant on AllPrep DNA spin column (Attention: carry no beats!)
9. Take filter and squeeze it
10. Centrifuge E-matrix-tube 1 min with full speed (14000 rpm)
11. Pipet rest of supernatant on spin column
12. Centrifuge column for 30 sec at 10000 rpm
13. **Flow-through for RNA-extraction** (*column for DNA-extraction*)
14. Add 1 volume of 70 % ethanol, mix it by pipetting
15. Transfer up to 700 µl of the sample to an RNeasy spin column (placed in a 2 ml collection tube) and centrifuge for 15 sec at 10000 rpm. Discard the flow-through.
16. **DNase digestion:**
    a) Add 350 µl RW1 buffer to RNeasy spin column, centrifuge for 15 sec at 10000 rpm and discard the flow-through.
    b) Add 10 µl DNase I (stock solution) to 70 µl RDD buffer, mix it by slightly inverting and shortly down-centrifugation.
    c) Place mixture on RNeasy spin column membrane, incubating at RT for 15 min
    d) Add 350 µl RW1 buffer, centrifuge for 15 sec at 10000 rpm and discard the flow-through.
17. Add 500 µl RPE buffer, centrifuge for 15 sec at 10000 rpm and discard the flow-through.
18. Add 500 µl RPE buffer, centrifuge for 2 min at 10000 rpm.
19. Place column in a new collection tube.
20. Centrifuge for 1 min with full speed
21. Place column in a 1.5 ml Eppi.
22. Add 30 µl RNase-free water, centrifuge for 1 min at 10000 rpm.
23. Place flow-through on column membrane and centrifuge for 1 min at 10000 rpm again.

Extraction of genomic DNA:

24. Add 500 µl AW1 buffer into AllPrep DNA spin column, centrifuge for 15 sec at 10000 rpm and discard the flow-through.
25. Add 500 µl AW2 buffer, centrifuge for 2 min with full speed.
26. Place column in 1.5 ml Eppi, add 50 µl EB buffer, incubate at RT for 1 min, centrifuge for 1 min at 10000 rpm.
27. Place flow-through on column membrane and centrifuge for 1 min at 10000 rpm again.

Supplemental Table 3: Tested chemical compositions to get fixed, intact cells of *S. salinarum* and the observed results after microscopically investigation of the cells.

tested fixative	chemical composition	observation
acidic Lugols solution	10 g KI, 5 g I_2, 95 ml H_2O_{dest}, 5 ml glacial acetic acid	cell shape intact and staining of the cell
alkaline Lugols solution	10 g KI, 5 g I_2, 100 ml H_2O_{dest}	cell shape intact and staining of the cell
Bouin`s solution	acetic acid (5 %), formaldehyde (9 %), picric acid (0,9 %)	cell lyses and bursting
Champy`s fixative	1,75 ml CrO_3 (1 %), 1,75 ml $K_2Cr_2O_7$ (3 %), 1 ml OsO_4 (2 %)	cell lyses and bursting
Da Fano's fixative	1% $CoNO_3$, 1% NaCl, 10% formalin	cell lyses and bursting
EtOH (100 %)	100 % EtOH	cell lyses and bursting
EtOH (70 %)	70 % EtOH	cell lyses and bursting
EtOH/Formalin solution	50 ml EtOH (70 %) + 5 ml formaldehyde (37 %)	cell lyses and bursting
formalin	45 ml H_2O_{dest} + 5 ml formaldehyde (37 %)	cell lyses and bursting
glutaraldehyde (1,5 %)	glutaraldehyde (1,5 %)	cell lyses and bursting
osmiumtetroxide	OsO_4	cell lyses and bursting
Stieve`s fluid	1,42 g $HgCl_2$ dissolved in 24 ml H_2O_{dest}, 3 ml formaldehyde, 1 ml acetic acid	cell lyses and bursting
tap water	tap water	cell lyses and bursting

Supplemental Table 4: Used samples and datasets of the study of Filker et al. (2017). These datasets comprised sequence data of protistan communities thriving between 39 and 440 ‰ salinities, originating from solar salterns located in France, Portugal, Spain, Chile, Argentina and on Cape Verde islands. Sample names are corresponding to the sample sames used in Fig. 14.

sample name	location	salinity [‰]
M43	Salinas d`Es Trenc, Mallorca, Spain	380
P39	Ria Formosa saltern, Faro, Portugal	380
SA8	Salinas de Boyeruca, Boyeruca, Chile	440
SA9	Salinas de Boyeruca, Boyeruca, Chile	440
SA6	Laguna del Diamante, Mendoza, Argentina	330
SP19	Salinas de Janubjo, Lanzarote, Spain	300
SP20	Salinas de Santa Pola, Santa Pola, Spain	310
SP17	Salines de S'Avall, Mallorca, Spain	300
SP15	Salinas de Ibiza, Ibiza, Spain	270
SP16	Campos, Mallorca, Spain	280
CV30	Pedra de Llume, Sal, Cap Verde Islands	300
CV27	Pedra de Llume, Sal, Cap Verde Islands	270
F28	Salin de Giraud, Arles, France	320
F27	Salin de Giraud, Arles, France	310
SP21	Salinas de la Trinitat, Delta de l'Ebre, Spain	290
SA11	Laguna Cejar (anthropogenic), San Pedro de Atacama, Chile	240
SA10	Laguna Cejar (pristine), San Pedro de Atacama, Chile	160
M41	Salinas d`Es Trenc, Mallorca, Spain	170
F26	Salin de Giraud, Arles, France	140
P37	Ria Formosa saltern, Faro, Portugal	120
M40	Salinas d`Es Trenc, Mallorca, Spain	90
P35	Ria Formosa saltern, Faro, Portugal	40
F25	Salin de Giraud, Arles, France	39

Supplemental Table 5: Sequence data overview listed after time points and replicates

replicate	number of raw reads	after trimming + Quality check	after SortMeRNA
control-1	62.959.518	50.888.532	33.454.930
control-2	58.731.700	47.775.976	31.833.139
control-3	69.817.402	56.519.212	34.975.358
2min-1	28.158.492	23.553.844	16.758.445
2min-2	29.261.476	25.338.108	19.247.584
2min-3	31.658.354	26.740.952	19.790.565
10min-1	33.727.842	28.996.324	22.089.689
10min-2	36.715.782	31.637.290	23.024.495
10min-3	35.570.716	31.443.136	22.633.375
30min-1	31.890.998	28.649.892	21.246.452
30min-2	34.269.518	30.950.130	23.171.617
30min-3	38.910.616	35.165.860	25.755.877
60min-1	32.654.766	28.695.268	20.969.535
60min-2	32.084.610	28.246.648	20.479.289
60min-3	35.399.852	31.071.208	22.720.934
120min-1	38.939.990	35.411.430	23.659.573
120min-2	30.191.302	26.581.652	17.811.227
120min-3	31.856.614	28.570.212	18.673.678
720min-1	42.292.670	38.190.118	24.506.878
720min-2	42.756.284	38.978.366	27.165.592
720min-3	43.292.202	39.730.758	28.248.755

Acknowledgements

First of all, I would like to thank Jun.-Prof. Dr. Sabine Filker and Prof. Dr. Thorsten Stoeck for the opportunity, support and independence needed to conduct this thesis. I would also thank both for their constructive feedback in the last four years. An additional thank appertain to Jun.-Prof. Dr. Sabine Filker for the sampling conducted around the world, especially in Mallorca, and for the possibility that I could be part of cruise AL491 with *RV Alkor*. Special thanks also go to Prof. Dr. Stefan Kins and Prof. Dr. Matthias Hahn for their willingness to act as chairman and second reviewer of my examination board.

I am deeply grateful to all members of the Department of Ecology and Molecular Ecology at the University of Kaiserslautern. You helped me a lot with your expertise, helpfulness and collegial support forming a relaxed lab environment. Special thanks in this case go to Hans-Werner Breiner, Larissa Frühe and our former members Dr. Guillaume Letendu and Dr. Lea Weinisch, who helped me a lot with trouble shooting and solving bioinformatic problems as well as mastering unexpected challenges. Additional thanks appertain to all the students involved in my PhD project, especially Carolin Schöbel. An equally thank goes to my so far unmentioned collaborators Prof. Dr. Antonio Pierik, Prof. Dr. Micah Dunthorn, Dr. Jan Dierking and the complete crew of AL491.

Furthermore, I would like to thank my friends in particular Felix Boos and especially Julian Laufer for their help by getting new ideas, for their motivation after uncounted setbacks and their exhortation that I am not getting lost in too many details missing the great connection by solving many bioinformatic issues. An equally thank goes to all persons, who helped me to improve this manuscript and were not mentioned so far.

The most important thank appertain to my family, namely Petra and Gerhard Mayer as well my sister Sarina Kühner, who supported me since I am able to think about. Without your help, patience and love this thesis would not be finished at all. Thanks for your motivation and the guaranteed freedom to me that I was able to pursue my goals and interests. Together we can master all challenges and strokes. Last but not least, I am thankful for the daily walks with my dog 'Timmy' and his patience and insistence because it helped me a lot to clear my mind and gave me the power to master all challenges occurring writing this thesis.

Curriculum vitae

Steffen Kühner

Master of Science

Education

11/2015 – 07/2019	**Ph.D. student**, Department of Molecular Ecology, University of Kaiserslautern
	Thesis: "Linking protistan community shifts along salinity gradients with cellular haloadaptation strategies"
	Thesis supervisor: Jun.-Prof. Dr. S. Filker
10/2012 - 07/2015	**M.Sc. (with honors)**, Department of Ecology, University of Kaiserslautern
	Thesis: "Glycine betaine is the main compatible solute in the halo-tolerant ciliates *Fabrea salina* and *Schmidingerothrix salinarum*" Thesis supervisor: Prof. Dr. T. Stoeck
10/2009 - 09/2012	**B.Sc.**, Department of Ecology, University of Kaiserslautern
	Thesis: "Zusammenhang zwischen der Abundanzverteilung intertidaler Sedimentbakterien und strukturierender Umweltfaktoren"
	Thesis supervisor: Prof. Dr. T. Stoeck
08/2000 - 03/2009	Higher Education Entrance Qualification Hohenstaufen-Gymnasium Kaiserslautern

Research & working experience

04/2017	Research cruise AL491 with *RV Alkor* (Protistan plankton community assemblages in the Eastern Baltic Sea)
since 11/2016	Teaching Assistant *Molecular Methods in Protistan Plankton Diversity Research*
08/2015	Research cruise M118 with *RV Meteor* (Latitudinal diversity gradient of protistan plankton from South England to the Cape Verde Islands)
11/2013 - 02/2014	Visiting student at Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS), Porto Alegre, Brasil
02/2012 - 03/2012	Internship at Senckenberg am Meer, Department of Marine Biology, Wilhelmshaven, Germany (Study of benthic organisms at the German Bight)

Awards & grants

07/2016	Award for best master thesis Kreissparkassenstiftung Kaiserslautern
06/2016	Award of the Society of Biochemistry and Molecular Biology (GBM) for an outstanding master thesis
06/2016	Award for the best master thesis in the Department of Biology, University of Kaiserslautern
11/2013	Research grant for Brasil, DAAD Promos (4 months)

Oral presentations

12/2018	"Microeukaryotic gene expression profiling along a pronounced salt gradient" Annual faculty meeting of the Department of Biology, University of Kaiserslautern, Altleiningen, Germany
02/2018	"Unravelling haloadaptation strategies of the heterotroph ciliate Schmidingerothrix salinarum using transcriptomics" 37. Congress of the German Society of Protozoologists, Cologne, Germany
12/2017	"Unravelling haloadaptation strategies of the heterotroph ciliate Schmidingerothrix salinarum using transcriptomics" Annual faculty meeting of the Department of Biology, University of Kaiserslautern, Thallichtenberg, Germany

Publications

Filker, S., **Kühner, S.**, Heckwolf, M., Dierking, J. and Stoeck, T. (2019) A fundamental difference between macrobiota and microbial eukaryotes: Protistan plankton has a species maximum in the freshwater-marine transition zone of the Baltic Sea. Environmental Microbiology, doi: 10.1111/1462-2920.14502

Weinisch, L., Kirchner, I., Grimm, M., **Kühner, S.**, Pierik, A., Rosselló-Móra, R. and Filker, S. (2018) Glycine Betaine and Ectoine Are the Major Compatible Solutes Used by Four Different Halophilic Heterotrophic Ciliates. Microbial Ecology, doi: 10.1007/s00248-018-1230-0

Weinisch, L., **Kühner, S.**, Roth, R., Grimm, M., Roth, T., Netz, D., Pierik, A. and Filker, S. (2018) Identification of osmoadaptive strategies in the halophile heterotrophic ciliate *Schmidingerothrix salinarum*. PLoS Biology, doi: 10.1371/journal.pbio.2003892

Kühner, S., Simao, T., Safi, L., Gazulha, F., Eizirik, E. and Utz, L. (2015) *Epistylis portoalegrensis* n. sp. (Ciliophora, Peritrichia): A New Freshwater Ciliate Species from Southern Brazil. Journal of Eukaryotic Microbiology, doi: 10.1111/jeu.12252-4653

Additional qualifications & skills

	German (mother tongue)
	English (fluently, C1 level)
	Spanish (advanced basic knowledge, B2 level)
	Portuguese (advanced basic knowledge, B2 level)
	Latin (Großes Latinum)
	Programming using Rstudio and Linux

12/2018	Workshop "Exploring the Company World" (University of Kaiserslautern)
11/2018	Workshop "Slidewriting - in 10 Minuten zum Publikumsmagneten" (University of Kaiserslautern)
10/2018	Seminar "Presentation Skills" (University of Kaiserslautern)
10/2018	Workshop "Getting it Done - Methoden und Motivation für den erfolgreichen Abschluss der Promotion" (University of Kaiserslautern)
09/2018	Seminar "In Führung gehen für Doktoranden" (University of Kaiserslautern)
08/2018	Seminar "Rhetorik, Auftreten und Präsentation im universitären Kontext" (University of Kaiserslautern)
08/2018	Workshop "Gespräche mit Mitarbeiter/innen/ Kollegen/innen zielorientiert führen" (University of Kaiserslautern)
01/2018	Workshop "Publizieren für Promovierende - Die Dissertation publizieren" (University of Kaiserslautern)
05/2017	Workshop "Scientific writing" (University of Kaiserslautern)
05/2017	Workshop "Peer-reviewed publishing" (University of Kaiserslautern)
01/2016	Workshop "Augmented Reality" (University of Kaiserslautern)
01/2016	Workshop "Motivation und Leichtigkeit im beruflichen Alltag" (University of Kaiserslautern)
12/2015	Workshop "Mastering the PhD" (University of Kaiserslautern)

Voluntary comittment & personal interests

03/2014 - 01/2019	honorary Youth leader of football department of Turn- und Sportgemeinde (TSG) 1861 Kaiserslautern
2009 - 2019	honorary football coach with DFB-Elite-Jugend-Lizenz (UEFA B-Level)
since 2007	Advanced Open Water Diver (AOWD) of SSI (130 logged dives)