PRIFYSGOL BANGOR / BANGOR UNIVERSITY

Diversity and community composition of pico- and nanoplanktonic protists in the Vistula River estuary (Gulf of Gdask, Baltic Sea)

Piwosz, Kasia; Calkiewicz, Joanna; Golebiewski, Marcin; Creer, Simon

Estuarine, Coastal and Shelf Science

DOI:
10.1016/j.ecss.2018.04.013

Published: 31/07/2018

Peer reviewed version

Cyswllt i'r cyhoeddiad / Link to publication

Dyfyniad o'r fersiwn a gyhoeddwyd / Citation for published version (APA):
Piwosz, K., Calkiewicz, J., Golebiewski, M., & Creer, S. (2018). Diversity and community composition of pico- and nanoplanktonic protists in the Vistula River estuary (Gulf of Gdask, Baltic Sea). *Estuarine, Coastal and Shelf Science*, *207*, 242-249. https://doi.org/10.1016/j.ecss.2018.04.013

Hawliau Cyffredinol / General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

11. May. 2021

Kasia Piwosz[1,2*], Joanna Całkiewicz[2], Marcin Gołębiewski[3,4], Simon Creer[5]

Diversity and community composition of pico- and nanoplanktonic protists in the Vistula River estuary (Gulf of Gdańsk, Baltic Sea).

Estuarine, Coastal and Shelf Science 207 (2018) 242–249

[1] Centre Algatech, Institute of Microbiology Czech Academy of Sciences, Novohradska 237, 37981 Třeboň, Czech Republic

[2] National Marine Fisheries Research Institute, ul. Kołłątaja 1, 81-332, Gdynia, Poland

[3] Chair of Plant Physiology and Biotechnology, Nicolaus Copernicus University, Lwowska 1, 87-100 Toruń, Poland

[4] Centre for Modern Interdisciplinary Research, Nicolaus Copernicus University, Wileńska 4, 87-100 Toruń, Poland

[5] School of Biological Sciences, Bangor University, Bangor, Gwynedd, LL57 2UW, UK

***Corresponding author:**

Kasia Piwosz, ORCID: 0000-0002-3248-3364

Postal address: Center Algatech, Institute of Microbiology Czech Academy of Sciences, Novohradska 237, 37981 Třeboň, Czech Republic

Telephone: +420 384 340 482

Fax: +420 384 340 415

E-mail: piwosz@alga.cz

E-mail addresses of other authors:

Joanna Całkiewicz: jcalkiewicz@mir.gdynia.pl

Marcin Gołębiewski: mgoleb@umk.pl

Simon Creer: s.creer@bangor.ac.uk

Abstract

Pico- and nanoplanktonic protists (eukaryotic microorganisms with cell size of < 3 µm and 3-20 µm, respectively) are the key component of plankton communities. However, their diversity and distribution patterns along environmental factors are still poorly recognized, largely due to their enormous phylogenetic diversity that has been realized only via the application of molecular methods over the past two decades. Here, we compared diversity and composition of active communities of pico- and nanoplantonic protists from three zones of the Vistula River estuary (Gulf of Gdańsk): freshwater, mixing (salinity 3.5) and brackish (salinity 7), in four seasons, by pyrosequencing the V3-V4 fragment of 18S rRNA taxonomy marker gene libraries. Alpha diversity was the highest at the brackish site, but the OTU richness was characteristic for specific protist groups. The active protistan communities in the freshwater and mixing zones (salinity 0-3.5) were similar (sharing >72% of phylotypes) and included centric diatoms (*Stephanodiscus minutulus*), synurophytes from clades C, E and F, and cryptophytes. However, at salinity of 7 at the brackish site the communities were significantly different from those in freshwater/mixing zone, and showed higher contributions of Dinophyceae, Mamiellophyceae, *Telonemia*, and picobiliphytes. The high similarity between the freshwater and mixing site, as well as high dissimilarity of the brackish site was observed in all months, despite seasonal shifts in pico- and nanoplantonic protistan communities. Seventy five percent of the observed variability in the communities was explained by combinations of temperature, salinity, nutrients and geographical distance, indicating interplay between species sorting and mass effects in shaping the active protistan communities in the Vistula River estuary. Groups that were more active in freshwaters and mixing zone seemed to be more affected by mass effects of mixing water masses, while those from brackish site by environmental species sorting. Finally, we report, for the first time, presence of Radiolaria (Acantharea) from the Baltic Sea.

Keywords

Picoplankton, nanoplankton, protists, estuary, Baltic Sea, diversity, distribution patterns

Introduction

Pico- and nanoplanktonic protists (eukaryotic microorganisms with cell size of < 3 µm and 3-20 µm, respectively) are the most abundant component of plankton communities (Lie et al., 2013; Piwosz et al., 2015a; Sherr et al., 2007). Due to small sizes and inconspicuous morphology, their diversity and distribution patterns remained unexplored until the application of molecular methods (Lopez-Garcia et al., 2001; Moon-van der Staay et al., 2001). Discovery of many hitherto unknown environmental groups (Guillou et al., 1999; Massana et al., 2002) has completely reshaped the eukaryotic tree of life (Adl et al., 2012; Burki, 2014; Hug et al., 2016). Still, their ecological and geographical patterns in space and time remains little explored and understood (de Vargas et al., 2015).

The Baltic Sea is among the largest brackish seas in the world. Salinity of the surface waters changes from 30 in Kattegat, where the water exchange with the North Sea occurs through a narrow and shallow Belt Sea, to < 1 in the Bothnian Bay, which is strongly influence by riverine run off. In the largest basin: Baltic Proper, salinity of the surface layer ranges between 7 and 8. Environmental gradients in the Baltic Sea strongly affects communities of microbial eukaryotes, and many typical marine groups, for example radiolarians or foraminifera, are absent (Hallfors, 2004). On the other hand, diversity of protists along these gradients seems to be unaffected, and peaks at horohalinicum (salinity 5-8 (Hu et al., 2016; Telesh et al., 2011). This unexpected large scale pattern in the open sea has been attributed to small cell size and rapid growth of planktonic protists, which allow them to rapidly adapt to new conditions (Telesh et al., 2013; Telesh et al., 2015). However, it remains to be seen whether similar diversity pattern occurs in coastal waters and river plumes, where activity of microorganisms, including protists, is much higher than in the open sea (Ameryk et al., 2005; Wasmund et al., 2001).

The Gulf of Gdańsk (Poland) lies on the southern Baltic Sea coast (Fig. 1). The run off of freshwaters from the Vistula River decreases its salinity compared to the Baltic Proper from 1 at the river mouth to about 6–7 in the open basin. The Vistula River also introduces freshwater protists into brackish environment, for example freshwater, aplastidic cryptophytes from lineage CRY1 (Piwosz et al., 2016). Moreover, changes in salinity affect protistan communities by promoting groups like pedinellids and MAST-6, and depressing phylotypes affiliated with typical marine groups, e.g. MALV-I alveolates (Piwosz and Pernthaler, 2010). Still, we lack detailed knowledge on how communities of pico- and nanoplanktonic protist changes from the river to the open waters of the Gulf of Gdańsk.

Here, we studied communities of active pico- and nanoplanktonic protists in the Vistula River (freshwater site, salinity (S) < 0.5), its plume (mixing zone, S~3.5) and brackish waters of the Gulf of Gdańsk (brackish site, S~7, Fig. 1), in four seasons, by high-throughput sequencing of V4 fragments of eukaryotic 18S rRNA amplified from environmental rRNA as a template. We provide insights into their diversity, distribution patterns, and environmental factors that can plausibly affect them in the coastal waters of the Baltic Sea.

Methods

Collection of samples

Triplicate samples of surface water were collected in July and October 2011, and in January and April 2012 along a salinity gradient from the Vistula River to the open waters of the Gulf of Gdańsk (36 samples in total, Fig. 1). Salinity and temperature were measured *in situ* with a Cast Away CTD probe (SonTec YSI Inc, USA).

Twenty-five litres of surface water were collected with a Niskin bottle. Twenty litres were filtered through a 20 µm mesh plankton net into acid and ethanol-sterilized canisters, washed thoroughly with the sampled water. These samples were used for RNA extraction and for cell

counts. Five litres of the unfiltered water were stored in light-proof canisters for analysis of nutrient and chlorophyll-*a*, and were processed immediately, as described below.

Nutrients

Subsamples of 0.5 litre of unfiltered water were collected into acid-clean containers and were stored at -20°C prior to downstream processing within a month of collection. Concentrations of total nitrogen (N-tot), $N\text{-}NO_3$, $N\text{-}NO_2$, and $N\text{-}NH_4$ (jointly referred to as dissolved inorganic nitrogen: DIN), total phosphorous (P-tot), soluble reactive phosphorous (SRP), and dissolved silicates (DSi) were determined according to Grasshoff et al. (1976).

Biological parameters

Concentrations of chlorophyll-*a* were measured in two fractions: total chlorophyll-*a* and chlorophyll-*a* < 20 µm (prefiltered first through a 20 µm plankton net). From each fraction, 10-50 ml were filtered onto glass-fiber GF/F filters (Whatmann) and stored at -20°C in the dark (< 1 month). Chlorophyll-*a* concentrations were measured using a fluorometric method after 24-hour extraction in 90% acetone in the dark at 4°C (Edler, 1979) with a Turner Designs 10-005R fluorometer.

For estimating abundance of heterotrophic pico- and nanoplankton, 2.5-50 ml of prefiltered water were filtered onto white polycarbonate filters (Cyclopore, Whatmann diameter 25 mm, pore size 0.8 µm). They were stained with 4',6-diamidino-2-phenylindole (DAPI, Sigma, concentration 5 µg/mL) for 10 minutes in the dark, mounted on microscope slides with Cargille oil A, and frozen at -20°C (Coleman, 1980). Samples were analysed by epifluorescent microscopy in UV light (Olympus BX50) under 1000× magnification. A minimum of 30 fields of views were analysed and at least 150 cells that did not show red chloroplasts fluorescence (i.e. heterotrophs) were counted.

Total RNA isolation

Pico- and nanoplankton biomass (fraction < 20 µm) from 2 litres of water was collected (using two filters per replicate) onto polycarbonate filters (0.4 µm pore size, 47 mm diameter, Cyclopore, Whatman, UK). The filtration time was < 30 min, and filters were immediately frozen at -80ºC and stored until RNA isolation the following day. Total RNA was extracted with GeneMATRIX Universal RNA Purification Kit according to the manufacturer's protocol including optional 10 minutes DNA digestion with DNaseI (Eurx, Gdańsk, Poland). The quality of the extracts and absence of genomic DNA were monitored with end-point PCR without the reverse transcription step, and agarose gel electrophoresis.

Reverse transcription, amplification of 18S rRNA fragments and sequencing

Reverse transcription was performed with a dART reverse transcriptase kit (Eurx) with TAReukREV3 reverse primer (Stoeck et al., 2010) at 45ºC for one hour. RNA was subsequently digested with an RNAse A (Eurx) for 30 minutes at 37ºC.

V4 18S rRNA fragments were amplified in a two-step PCR process (Schülke 2000), with TAReuk454FWD1 and TAReukREV3 primers (Stoeck et al., 2010) using the high fidelity Pfu polymerase. Amplicons were purified with a GeneMATRIX Agarose-out DNA Purification Kit (Eurx) after the first PCR and using Qiaquick Gel Extraction Kit (Qiagen) following the second PCR. Concentrations of purified DNA fragments were measured with PicoGreen kit (LifeTechnologies, Molecular Probes) on a Perkin-Elmer LS-5B fluorometer. Eighteen samples were pooled in equimolar amounts to a final concentration > 10 ng μl^{-1}, and sequenced on 454 FLX Titanium platform (Centre for Genomic Research, University of Liverpool, UK).

Bioinformatic analyses

Bioinformatic analyses followed the standard operating procedure of the Schloss group (www.mothur.org/wiki/454_SOP) and were performed in Mothur v.1.32 (Schloss et al., 2009). The Schloss procedure was modified with custom-tailored Perl scripts to improve denoising and chimera removal, and also to produce the list of shared OTUs from averaged 10 subsamples of the whole data. Below we provide a short summary of the key steps, and the detailed procedure is described in the Supplementary File 1.

The flows were extracted separately for forward and reverse reads, and they were assigned to the samples base on the barcode sequences. We used long barcodes (10 nt) that differed by at least four nucleotides (Hamming and Levenshtein distances = 4) to minimize incorrect assignments (Faircloth and Glenn, 2012).

Demultiplexed sequences were trimmed to 500-650 flows and denoised with AmpliconNoise algorithms. The sequencing and PCR noise was removed with Single Linkage pre-clustering (Huse et al., 2010). Chimera removal was performed in three rounds: i) with UCHIME in *de novo* model (Edgar et al., 2011), ii) with PERSEUS (Quince et al., 2011) and iii) with chimera slayer (Haas et al., 2011) using the PR2 template alignment (Guillou et al., 2013).

Full-length sequences were used for classification with a naive Bayesian classifier (Wang et al., 2007) with the PR2 template and taxonomy files (downloaded at http://ssu-rrna.org/pr2 on May 14, 2014, (Guillou et al., 2013)) at the bootstrap confidence level of 80%. Average linkage (UPGMA) algorithm was used to construct OTUs at the 0.03 dissimilarity level. Singletons, doubletons and taxa assigned as 'unknown' were removed from the data.

To ensure that OTUs frequencies in the subsampled dataset are close to the original ones, the final reads set was subsampled ten times to 2500 reads per sample, subsamples were combined, the whole set was dereplicated and used for distance matrix calculation and OTU

construction via average neighbour clustering at 97% similarity level. A shared OTU table was constructed, and the table averaged over the subsamples (i.e. for each OTU numbers of reads found in each subsample were summed and the sum was divided by ten) was calculated with a Perl script. OTUs were classified using consensus approach with PR2 taxonomic assignment (Guillou et al., 2013).

We estimated sequencing error rate by processing the V4 fragment of 18S rRNA gene of *Skeletonema marinoi* BA98 from the Culture Collection of Baltic Algae (University of Gdańsk) (Pniewski et al., 2010) amplified from genomic DNA isolated from pure culture with the seq.error command of mothur. The PR2 database (Guillou et al., 2013) was used as the reference templates set for ChimeraSlayer. The *S. marinoi* BA98 18S rRNA gene sequence (HM805045.1) was used as a reference.

Statistical analyses

Statistical analysis were performed in Primer 7 (Clarke et al., 2014) with the PERMANOVA add on package (Anderson et al., 2008). Differences between the environmental conditions at the sites were analysed with principal component analysis (PCA), test of homogeneity of dispersions (PERMDISP) and Permutational ANOVA and MANOVA (PERMANOVA). Pico- and nanoplankton communities were analysed with PERMDISP and PERMANOVA. Relationships between environmental variables and community composition were explored with distance based Redundancy Analysis (dbRDA). Correlations between the environmental variables and community composition, and geographical distance and community composition were done by Mantel test. Prior to these analyses, the environmental data were log(X+1) transformed and normalized, and the community data were reduce to presence/absence matrix to account for low quantitative accuracy of the amplicon data.

Rarefaction curves were calculated in R (R Core Team 2015) package vegan ver. 2.3-4 (Oksanen et al., 2018) with function *rarecurve*.

Results

Environmental conditions

Environmental conditions significantly differed between all sites (PERMANOVA, $p<0.05$). In general, the freshwater site was most eutrophic and the brackish site was most oligotrophic (Supplementary File 2). Nutrient concentrations in the mixing zone corresponded to the shares of fresh- and brackish waters, except for summer, when they were significantly lower than expected (Supplementary Table S1). Seasonally, concentrations of nutrients were always the highest in January, and the lowest in July. The range of values of the environmental factors was similar for the freshwater site and the mixing zone (PERMDISP analysis, $p<0.41$), and it was significantly larger than at the brackish site (PERMDISP analysis, $p<0.01$), indicating that environmental conditions varied more in the Vistula River and the mixing zone than in the open waters of the Gulf of Gdańsk. The Principal Component Analysis of the environmental variables clustered the samples according to site and season (Fig. 2). The first principal component that explained 51.6 % of the variance, correlated positively with salinity and temperature, and negatively with concentrations of dissolved silica, total phosphorus and total nitrogen. The second principal component explained 37.8 % of the variance, and correlated positively with chlorophyll-*a* < 20 µm, and abundance of heterotrophic pico- and nanoplankton, and negatively with salinity.

Pico- and nanophytoplankton were the key component of phytoplankton community: chlorophyll-*a* < 20 µm contributed from 54 to 99% (85±12% on average) of total chlorophyll-*a* (Supplementary File 2). The concentrations of chlorophyll-*a* in both fractions were up to 25-fold higher at the freshwater site than at the brackish site during the maxima in April 2012

(respectively, total and < 20 µm fractions freshwater site: 42.5 µg L^{-1} and 41.9 µg L^{-1}, mixing zone: 31.9 µg L^{-1} and 29.5 µg L^{-1}, brackish site: 2.2 µg L^{-1} and 1.6 µg L^{-1}), and in July 2011 (freshwater site: 77.0 µg L^{-1} and 60.3 µg L^{-1}, mixing zone: 49.8 µg L^{-1} and 27.1 µg L^{-1}, brackish site: 3.8 µg L^{-1} and 3.2 µg L^{-1}). The abundance of heterotrophic pico- and nanoplankton ranged from 400±100 to 16 400 ±2 700 cell mL^{-1}, and corresponded with the spatial and temporal patterns observed for chlorophyll-*a* (Supplementary File 2).

Sequencing statistics

A total of 885 380 (217 726 unique) raw reads were generated and denoising left 95 675 unique sequences. 754 169 sequences (53 442 unique) covered the target region of the SILVA alignment (13 900-22 400). Chimeric sequences were identified by a strict, three-step procedure: UCHIME removed 29 161 (13 542 unique) sequences, Perseus 11 036 (1684 unique), and chimera.slayer 16 842 (786 unique). 697 130 sequences were left, all of which were affiliated with the Eukaryota domain. Upon culling singletons and doubletons, 693 600 (9 750 unique) sequences were used for downstream analyses. The reads have been deposited in the NCBI Sequence Read Archive database under accession number SRP096863. The error rate was estimated to be 6.28×10^{-5} errors/base.

Active pico –and nanoplankton communities in the Vistula River estuary

A total of 1237 OTUs were observed at the 97% similarity level (Supplementary Table S2). Species accumulation and rarefaction curves started to plateau but did not reach a clear asymptote, indicating that the diversity of the whole estuary was moderately sampled (Fig. 3A).

The number of observed OTUs was similar at all sites (Fig. 3B), but the diversity indices Shannon and Pielou's evenness, were significantly higher at the brackish site (Fig. 3C and D,

ANOVA: $p<0.001$, post-hoc Holm-Sidak pairwise comparison: $p<0.001$). This indicates a more even distribution of OTUs in the Gulf of Gdańsk, and lack of a clearly dominant phylotype. Seasonally, the significantly lower values of alpha diversity indices were in October at all sites (ANOVA: $p<0.001$, post-hoc Holm-Sidak pairwise comparison: $p<0.001$, Fig.3).

The OTU richness of different taxonomic groups varied between the zones of the estuary (Table 1). For instance, within the Alveolata, ciliates were more diverse in the mixing zone, while dinoflagellates exhibited higher diversity at the brackish site. Distinct diversity patterns were observed also at lower taxonomical levels, for example within ciliates: Oligohymenophorea and Prostomatea had the highest number of OTUs in the mixing zone; the diversity of Litostomatea increased from the freshwater to the brackish site, while Spirotrichea had a similar number of OTUs in all zones. Differences OTUs richness in specific zones of estuary occurred also in other taxonomic groups (Table 1).

From the all OTUs detected in the Vistula River estuary, only 148 (12%) occurred at all sites (Fig. 4A). The communities of pico- and nanoplankton in the freshwater and mixing zone shared over 450 of OTUs and were very similar (Fig. 4A, PERMANOVA, $p=0.5$, average Bray-Curtis similarity between the samples: 45.2%). They were dominated by centric diatoms (Mediophyceae, Coscinodiscophyceae) (Fig. 4B): *Stephanodiscus minutulus* contributed substantially to the diatom reads in July and October (comprising over 75% of the reads in the October libraries), and co-dominated with *Cyclotella* sp. and *Skeletonema* sp. in January and April (Supplementary Table S2). Contributions of reads from other groups varied between the seasons (Fig. 4B). In June and April, centric diatoms were accompanied by Synurophyceae from clades C, E and F, *Cryptomonas* sp. and basal cryptophytes from heterotrophic CRY1 lineage, in January in addition to centric diatoms there was a higher

contribution of sequences from synurophytes and heterotrophic groups like cercozoa and ciliates (Fig. 4B, Supplementary Table S2).

Active pico- and nanophytoplankton communities in the brackish waters of the Gulf of Gdańsk shared only 257 OTUs with those in the freshwater and mixing zone (Fig. 4A), and differed from them significantly (PERMANOVA, p=0.001, average Bray-Curtis dissimilarity between the samples in the brackish zone and freshwater-mixing zone: 89.8%). Reads from dinoflagellates and ciliates (Alveolates) were abundant in the brackish samples, but unlike in the freshwater and mixing zones, a clearly dominant group was lacking (Fig. 4B, Supplementary Table S2). The characteristic groups for the brackish waters included Mamiellophyceae, *Telonemia*, Picobiliphyta, Dictyochophyceae, Choanoflagellida, Spirotrichea, and Dinophyceae (Fig. 5A), whose contribution to communities in July and April was similar (Figs. 4B). Centric diatoms: *S. minutulus* and *Skeletonema* sp., and Dinophyceae increase their contribution in October, while in January higher contributions from ebriids (Thecofilosea), haptophytes (Prymnesiales, Pavlovophyceae) and *Pyramimonas* sp. were detected (Fig. 4B, Supplementary Table S2).

Correlations with environmental variables

Temperature, salinity and nutrients collectively explained 75.6 % of total observed variability in the pico- and nanoplankton communities, with the first two axes explaining 71.3% of the total variation (dbRDA, $p<0.05$, Fig. 5B). The first dbRDA axis correlated positively with salinity and negatively with dissolved silica and total phosphorous and the second dbRDA axis correlated positively with temperature.

The community composition of pico- and nanoplankton was strongly associated to the environmental variables and the geographical distance between the samples. The strength of these associations was similar except for in summer, when the correlation with geographical

distance was stronger (Table 2). However, it differed for specific protistan groups. For instance, centric diatoms (characteristic for the freshwater site and mixing zone) were more strongly correlated with geographical distance, while Mamiellophyceae (characteristic for the brackish site) with environmental parameters (Table 3).

Discussion

In this study, we contributed to the knowledge on pico- and nanoplankton diversity and community composition in the coastal waters of the Baltic Sea. The microplanktonic communities (unicellular organisms with cell size > 20 µm) have been well studied by light microscopy in both in the open Baltic Sea (Feuerpfeil et al., 2004; Gasiunaite et al., 2005; Olenina et al., 2006; Suikkanen et al., 2007; Telesh et al., 2011; Wasmund et al., 2017), and the Gulf of Gdańsk for many years (Kownacka et al., 2013; Wielgat-Rychert et al., 2013; Witek et al., 1997a). Unfortunately, pico- and nanoplanktonic cells cannot be easily recognized by light microscopy (Piwosz et al., 2016), and their molecular studies from the Baltic Sea are still rare (Majaneva et al., 2012; Piwosz and Pernthaler, 2010, 2011; Piwosz et al., 2015b). Our study is among the first that exhaustively described pico- and nanoplankton communities in an estuary of the brackish Baltic Sea using a high throughput sequencing method. To our knowledge, this is the first account of radiolarians from the Baltic Sea (Hallfors, 2004; Hu et al., 2016), and of pelagophytes, amoebozoans and apusozoans from the Vistula River and the Gulf of Gdańsk (Piwosz and Pernthaler, 2010, 2011; Rychert et al., 2013).

We used amplicons of 18S rRNA generated directly from extracted RNA, to focus on active protistan communities in the Vistula river estuary, because we were concerned that presence of DNA from dead cells would create a misleading picture of survival of phylotypes in different zones of estuary. A focus on rRNA was also the main reason why we did not use

rRNA:rDNA ratios as a proxy for protist activity, in addition to other limitations of this approach (Blazewicz et al., 2013). A number of reads originating from a specific phylotype in the libraries generated here might have resulted either from changes in its activity, or changes in its abundance in the different zones of the estuary. Thus, it is not a direct measure of activity. To overcome this hindrance, we avoided comparison of the abundance of reads between different OTUs, but instead we compared the abundance of reads of specific OTUs between the samples (Gołębiewski et al., 2017; Ibarbalz et al., 2014).

Patterns of pico- and nanoplankton protists diversity in the Vistula estuary

The alpha diversity of pico- and nanoplankton protists in the Vistula River estuary was the highest at the brackish site, as indicated by values of Shannon diversity and Pielou's evenness indices (Fig. 3C, D). This seems to agree with the large scale pattern observed for the whole Baltic Sea, where the number of taxa of planktonic protists also peaks at salinities between 5-8 (horohalinicum) (Telesh et al., 2013; Telesh et al., 2011), although we lack data from higher salinities and open Baltic Proper. Interestingly, the diversity patterns observed here for pico- and nanoplanktonic protists were very different from those observed for bacteria in the Vistula estuary (Gołębiewski et al., 2017). Bacterial and protistan diversity patterns also differed at the scale of the whole Baltic Sea salinity gradient (Herlemann et al., 2011; Telesh et al., 2015).

The Vistula River has a pronounced effect on microbial processes in the Gulf of Gdańsk (Ameryk et al., 2005; Wielgat-Rychert et al., 2013; Witek et al., 1997b). It also contributed many pico- and nanoplanktonic protistan phylotypes to the Gulf of Gdańsk (Fig. 4), as previously observed for phytoplankton and bacteria (Gołębiewski et al., 2017; Wielgat-Rychert et al., 2013). Nevertheless, only few phylotypes were common for the whole estuary (Fig. 4A), and the active pico- and nanoplanktonic communities differed significantly

between the brackish waters of the Gulf of Gdańsk and less saline waters of the mixing zone and the Vistula River (Fig. 5). Similar patterns in protists distribution along the increasing salinity, e.g. the replacement of diatoms with dinoflagellates and cryptophytes, was also observed in other estuaries (Balzano et al., 2015; Bazin et al., 2014a; Bazin et al., 2014b; Herfort et al., 2011; Lee et al., 2017). The observed differences between the sites were significant in all the investigated season (Figs. 4 and 5). The temporal resolution of our study was low, but still higher than in most study that usually focuses on summer season (Hu et al., 2016; Wielgat-Rychert et al., 2013). The dynamics of planktonic protist is high in the Gulf of Gdańsk (Kownacka et al., 2013; Piwosz and Pernthaler, 2010; Piwosz et al., 2015b), but it is plausible that similar patterns in beta diversity of pico- and nanoplanktonic protists in the Vistula River estuary can be observed most of the time.

The correlations between the community composition, environmental factors, and geographical distance were very strong (Table 2, Fig. 5A), indicating similar importance of species sorting by environmental factors, and mass effects from mixing of different water masses on the distribution of pico- and nanoplanktonic protists (Lallias et al., 2015; Lindstrom and Langenheder, 2012). The strength of these correlations, however, differed for specific groups (Table 3). Groups that were more represented at the freshwater site were correlated stronger with the distance, indicating the dilution effect due to mixing of freshwater and brackish water masses (Wielgat-Rychert et al., 2013). In contrast, marine groups seem to have been more affected by environmental factors, mostly salinity (Fig. 5). Indeed, it had been previously observed that even slight change in salinity may cause pronounced changes in abundance of some pico- and nanoplanktonic protists in the Gulf of Gdańsk (Piwosz and Pernthaler, 2010). On the other hand, we did not investigate food webs factors, like grazing by meso- and microzooplankton, or bacterial food availability for bacterivorous hetero- and mixotrophic protists (Piwosz and Pernthaler, 2011; Rychert, 2016; Witek et al., 1997a), which

likely are important considering elevated microbial activity in the Vistula River plumes (Ameryk et al., 2005; Wielgat-Rychert et al., 2013; Witek et al., 1997b). Further such ecological network studies would be important for explaining processes shaping protistan communities in estuaries, which are places of pivotal importance for understanding ecological and biogeochemical processes in coastal zones (Lunau et al., 2013; Schiewer and Schernewski, 2004).

Conclusions

With this study, we contributed to knowledge of spatial distribution patterns of pico- and nanoplanktonic protists by describing active communities along an ecological gradient in a brackish estuary. We report, to our knowledge for the first time, presence of pelagophytes, amoebozoans and apusozoans from the Vistula River and the Gulf of Gdańsk, and of radiolarians from the Baltic Sea. Our main conclusions are:

- Communities of pico- and nanoplanktonic protists were similar in the freshwater Vistula River and its mixing zone, and differed from those in the brackish waters of the Gulf of Gdańsk;
- Diversity of pico- and nanoplanktonic protists was the highest at the brackish site, which agrees with the large scale macroecological pattern observed for the whole Baltic Sea;
- The species sorting and mass effects seems to have been of similar importance in shaping the composition of communities pico- and nanoplanktonic protists in the Vistula River estuary;
- The distribution of freshwater groups in the Gulf of Gdańsk might have resulted mainly from mass effects, while marine groups present in the estuary are likely to be more affected by species sorting.

Acknowledgments

We thank Jerzy Lissowski, the cpt. of m/y Hestia, Bartosz Witalis, Tycjan Wodzinowski and Robert Gąsiorek for help with sampling, Lena Szymanek for the map of the Vistula estuary (Fig. 1), Michał Piwosz for calculating distances between the sampling stations, Mariusz Zalewski for measuring concentrations of chlorophyll-a, Hanna Wróblewska and Marianna Pastuszak for measuring concentrations of nutrients, and Katarzyna Spich for microscopy

This study was supported by project 795/N-CBOL/2010/0 founded by the National Science Centre Poland and project DEMONA (PSPB-036/2010) supported by a grant from Switzerland through the Swiss contribution to the enlarged European Union (Polish-Swiss Research Programme), and by the Czech Ministry of Education projects Algatech Plus (LO1416). Authors' contribution: , KP – study design, sampling, interpretation of the results, writing the manuscript, MG – study design, bioinformatics, critical appraisal of the manuscript, JC – sampling, laboratory molecular work, critical appraisal of the manuscript, SC – study design, bioinformatics, critical appraisal of the manuscript.

References

Adl, S.M., Simpson, A.G.B., Lane, C.E., Lukes, J., Bass, D., Bowser, S.S., Brown, M.W., Burki, F., Dunthorn, M., Hampl, V., Heiss, A., Hoppenrath, M., Lara, E., le Gall, L., Lynn, D.H., McManus, H., Mitchell, E.A.D., Mozley-Stanridge, S.E., Parfrey, L.W., Pawlowski, J., Rueckert, S., Shadwick, L., Schoch, C.L., Smirnov, A., Spiegel, F.W., 2012. The Revised Classification of Eukaryotes. J. Eukaryot. Microbiol. 59, 429-493.

Ameryk, A., Podgorska, B., Witek, Z., 2005. The dependence between bacterial production and environmental conditions in the Gulf of Gdansk. Oceanologia 47, 27-45.

Anderson, M.R., Gorley, R., Clarke, K.R., 2008. PERMANOVA+ for PRIMER: Guide to Software and Statistical Methods. PRIMER-E, Plumouth, UK.

Balzano, S., Abs, E., Leterme, S.C., 2015. Protist diversity along a salinity gradient in a coastal lagoon. Aquat. Microb. Ecol. 74, 263-277.

Bazin, P., Jouenne, F., Deton-Cabanillas, A.-F., Perez-Ruzafa, A., Veron, B., 2014a. Complex patterns in phytoplankton and microeukaryote diversity along the estuarine continuum. Hydrobiologia 726, 155-178.

Bazin, P., Jouenne, F., Friedl, T., Deton-Cabanillas, A.F., Le Roy, B., Veron, B., 2014b. Phytoplankton Diversity and Community Composition along the Estuarine Gradient of a Temperate Macrotidal Ecosystem: Combined Morphological and Molecular Approaches. Plos One 9, e94110.

Blazewicz, S.J., Barnard, R.L., Daly, R.A., Firestone, M.K., 2013. Evaluating rRNA as an indicator of microbial activity in environmental communities: limitations and uses. Isme J. 7, 2061-2068.

Burki, F., 2014. The Eukaryotic Tree of Life from a Global Phylogenomic Perspective. Cold Spring Harb Perspect Biol 6, a016147.

Clarke, K., Gorley, R., Somerfield, P., Warwick, R., 2014. Change in marine communities: an approach to statistical analysis and interpretation, 3rd edition ed. PRIMER-E, Plymouth, UK.

440 Coleman, A.W., 1980. Enhanced detection of bacteria in natural environments by
441 fluorochrone staining of DNA. Limnol. Oceanogr. 25, 948-951.

442 de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahe, F., Logares, R., Lara, E., Berney, C.,
443 Le Bescot, N., Probert, I., Carmichael, M., Poulain, J., Romac, S., Colin, S., Aury, J.M.,
444 Bittner, L., Chaffron, S., Dunthorn, M., Engelen, S., Flegontova, O., Guidi, L., Horak, A.,
445 Jaillon, O., Lima-Mendez, G., Lukes, J., Malviya, S., Morard, R., Mulot, M., Scalco, E.,
446 Siano, R., Vincent, F., Zingone, A., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S.,
447 Acinas, S.G., Bork, P., Bowler, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Not,
448 F., Ogata, H., Pesant, S., Raes, J., Sieracki, M.E., Speich, S., Stemmann, L., Sunagawa, S.,
449 Weissenbach, J., Wincker, P., Karsenti, E., Tara Oceans, C., 2015. Eukaryotic plankton
450 diversity in the sunlit ocean. Science 348.

451 Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R., 2011. UCHIME improves
452 sensitivity and speed of chimera detection. Bioinformatics 27, 2194-2200.

453 Edler, L., 1979. Recommendations on methods for marine biological studies in the Baltic Sea.
454 Phytoplankton and Chlorophyll. Baltic Marine Biologist Publication 5, 1-38.

455 Faircloth, B.C., Glenn, T.C., 2012. Not All Sequence Tags Are Created Equal: Designing and
456 Validating Sequence Identification Tags Robust to Indels. PLOS ONE 7, e42543.

457 Feuerpfeil, P., Rieling, T., Estrum-Youseff, S.R., Dehmlow, J., Papenfuß, T., Schoor, A.,
458 Schiewer, U., Schubert, H., 2004. Carbon budget and pelagic community compositions at two
459 coastal areas that differ in their degree of eutrophication, in the Southern Baltic Sea. Estuar.
460 Coast. Shelf Sci. 61, 89-100.

461 Gasiunaite, Z.R., Cardoso, A.C., Heiskanen, A.S., Henriksen, P., Kauppila, P., Olenina, I.,
462 Pilkaityte, R., Purina, I., Razinkovas, A., Sagert, S., Schubert, H., Wasmund, N., 2005.
463 Seasonality of coastal phytoplankton in the Baltic Sea: Influence of salinity and
464 eutrophication. Estuar. Coast. Shelf Sci. 65, 239-252.

465 Gołębiewski, M., Całkiewicz, J., Creer, S., Piwosz, K., 2017. Tideless estuaries in brackish
466 seas as a possible freshwater-marine transition zones for bacteria – the case study of the
467 Vistula river estuary. Environmental Microbiology Reports 9, 129–143.

468 Grasshoff, K., Ehrhardt, M., Kremling, K., 1976. Methods for sea water analysis. Verlag
469 Chemie, 1-419.

470 Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., Boutte, C., Burgaud, G.,
471 de Vargas, C., Decelle, J., del Campo, J., Dolan, J.R., Dunthorn, M., Edvardsen, B.,
472 Holzmann, M., Kooistra, W., Lara, E., Le Bescot, N., Logares, R., Mahe, F., Massana, R.,
473 Montresor, M., Morard, R., Not, F., Pawlowski, J., Probert, I., Sauvadet, A.L., Siano, R.,
474 Stoeck, T., Vaulot, D., Zimmermann, P., Christen, R., 2013. The Protist Ribosomal Reference
475 database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with
476 curated taxonomy. Nucleic Acids Res. 41, D597-D604.

477 Guillou, L., Moon-van der Staay, S.-Y., Claustre, H., Partensky, F., Vaulot, D., 1999.
478 Diversity and abundance of Bolidophyceae (Heterokonta) in two oceanic regions. Appl.
479 Environ. Microbiol. 65, 4528-4536.

480 Haas, B.J., Gevers, D., Earl, A.M., Feldgarden, M., Ward, D.V., Giannoukos, G., Ciulla, D.,
481 Tabbaa, D., Highlander, S.K., Sodergren, E., Methe, B., DeSantis, T.Z., Petrosino, J.F.,

Knight, R., Birren, B.W., Human Microbiome, C., 2011. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. Genome Res. 21, 494-504.

Hallfors, G., 2004. Checklist of Baltic Sea Phytoplankton Species (including some heterotrophic protistan groups). Helsinki Commission, Helsinki.

Herfort, L., Peterson, T.D., McCue, L.A., Zuber, P., 2011. Protist 18S rRNA gene sequence analysis reveals multiple sources of organic matter contributing to turbidity maxima of the Columbia River estuary. Mar. Ecol. Prog. Ser. 438, 19-31.

Herlemann, D.P.R., Labrenz, M., Juergens, K., Bertilsson, S., Waniek, J.J., Andersson, A.F., 2011. Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. Isme J. 5, 1571-1579.

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.

Hug, L.A., Baker, B.J., Anantharaman, K., Brown, C.T., Probst, A.J., Castelle, C.J., Butterfield, C.N., Hernsdorf, A.W., Amano, Y., Ise, K., Suzuki, Y., Dudek, N., Relman, D.A., Finstad, K.M., Amundson, R., Thomas, B.C., Banfield, J.F., 2016. A new view of the tree of life. Nature Microbiology 1, 16048.

Huse, S.M., Welch, D.M., Morrison, H.G., Sogin, M.L., 2010. Ironing out the wrinkles in the rare biosphere through improved OTU clustering. Environ. Microbiol. 12, 1889-1898.

Ibarbalz, F.M., Perez, M.V., Figuerola, E.L.M., Erijman, L., 2014. The Bias Associated with Amplicon Sequencing Does Not Affect the Quantitative Assessment of Bacterial Community Dynamics. Plos One 9, e99722.

Kownacka, J., Edler, L., Gromisz, S., Łotocka, M., Olenina, I., Ostrowska, M., Piwosz, K., 2013. Non-indigenous species Chaetoceros cf. lorenzianus Grunow 1863 – A new, predominant component of autumn phytoplankton in the southern Baltic Sea. Estuar. Coast. Shelf Sci. 119, 101-111.

Lallias, D., Hiddink, J.G., Fonseca, V.G., Gaspar, J.M., Sung, W., Neill, S.P., Barnes, N., Ferrero, T., Hall, N., Lambshead, P.J.D., Packer, M., Thomas, W.K., Creer, S., 2015. Environmental metabarcoding reveals heterogeneous drivers of microbial eukaryote diversity in contrasting estuarine ecosystems. Isme J. 9, 1208-1221.

Lee, J.E., Chung, I.K., Lee, S.R., 2017. Dynamic genetic features of eukaryotic plankton diversity in the Nakdong River estuary of Korea. Chin. J. Oceanol. Limnol. 35, 844-857.

Lie, A.A.Y., Kim, D.Y., Schnetzer, A., Caron, D.A., 2013. Small-scale temporal and spatial variations in protistan community composition at the San Pedro Ocean Time-series station off the coast of southern California. Aquat. Microb. Ecol. 70, 93-110.

Lindstrom, E.S., Langenheder, S., 2012. Local and regional factors influencing bacterial community assembly. Environmental Microbiology Reports 4, 1-9.

Lopez-Garcia, P., Rodriguez-Valera, F., Pedros-Alio, C., Moreira, D., 2001. Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. Nature 409, 603-607.

Lunau, M., Voss, M., Erickson, M., Dziallas, C., Casciotti, K., Ducklow, H., 2013. Excess nitrate loads to coastal waters reduces nitrate removal efficiency: mechanism and implications for coastal eutrophication. Environ. Microbiol. 15, 1492-1504.

Majaneva, M., Rintala, J.-M., Hajdu, S., Hallfors, S., Hallfors, G., Skjevik, A.-T., Gromisz, S., Kownacka, J., Busch, S., Blomster, J., 2012. The extensive bloom of alternate-stage *Prymnesium polylepis* (Haptophyta) in the Baltic Sea during autumn-spring 2007-2008. Eur. J. Phycol. 47, 310-320.

Massana, R., Guillou, L., Diez, B., Pedros-Alio, C., 2002. Unveiling the Organisms behind Novel Eukaryotic Ribosomal DNA Sequences from the Ocean. Appl. Environ. Microbiol. 68, 4554-4558.

Moon-van der Staay, S.Y., De Wachter, R., Vaulot, D., 2001. Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. Nature 409, 607-610.

Oksanen, J., Blanchet, F., Kindt, R., Legendre, P., Minchin, P., O'Hara, R., Simpson, G., Solymos, P., Henry, M., Stevens, H., Wagner , H., 2018. vegan: Community Ecology Package. R package version 2.4-6, https://CRAN.R-project.org/package=vegan.

Olenina, I., Hajdu, S., Edler, L., Andersson, A., Wasmund, N., Busch, S., Gobel, J., Gromisz, S., Huseby, S., Huttunen, M., Jaanus, A., Kekkonen, P., Ledaine, I., Ziemkiewicz, E., 2006. Biovolumes and size-classes of phytoplankton in the Baltic Sea.

Piwosz, K., Kownacka, J., Ameryk, A., Zalewski, M., Pernthaler, J., 2016. Phenology of cryptomonads and the CRY1 lineage in a coastal brackish lagoon (Vistula Lagoon, Baltic Sea). J. Phycol. 52, 626–637

Piwosz, K., Pernthaler, J., 2010. Seasonal population dynamics and trophic role of planktonic nanoflagellates in coastal surface waters of the Southern Baltic Sea. Environ. Microbiol. 12, 364-377.

Piwosz, K., Pernthaler, J., 2011. Enrichment of Omnivorous Cercozoan Nanoflagellates from Coastal Baltic Sea Waters. PLoS ONE 6, e24415.

Piwosz, K., Spich, K., Całkiewicz, J., Weydmann, A., Kubiszyn, A.M., Wiktor, J.M., 2015a. Distribution of small phytoflagellates along an Arctic fjord transect. Environ. Microbiol. 17, 2393–2406.

Piwosz, K., Villiger, J., Pernthaler, J., 2015b. Numbers vs. reads: Comparison of patterns in algal dynamics revealed by high throughput sequencing and microscopic counts. Eur. J. Phycol. 50, 134-134.

Pniewski, F., Friedl, T., Latała, A., 2010. Identification of diatom isolates from the Gulf of Gdańsk: Testing of species identifications using morphology, 18S rDNA sequencing and DNA barcodes of strains from the Culture Collection of Baltic Algae (CCBA). Oceanol. Hydrobiol. Stud. 39, 3-20.

Quince, C., Lanzen, A., Davenport, R.J., Turnbaugh, P.J., 2011. Removing Noise From Pyrosequenced Amplicons. BMC Bioinformatics 12, 38.

R Core Team 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org/.

Rychert, K., 2016. Growth rates of common pelagic ciliates in a highly eutrophic lake measured with a modified dilution method. Oceanol. Hydrobiol. Stud. 45, 216-229.

Rychert, K., Spich, K., Laskus, K., Paczkowska, M., Wielgat-Rychert, M., Sojda, G., 2013. Composition of protozoan communities at two sites in the coastal zone of the southern Baltic Sea. Oceanol. Hydrobiol. Stud. 42, 268-276.

Schiewer, U., Schernewski, G., 2004. Self-purification capacity and management of Baltic coastal ecosystems. J. Coast. Conservation 10, 25-32.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., Weber, C.F., 2009. Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Appl. Environ. Microbiol. 75, 7537-7541.

Schülke , M., 2000. An economic method for the fluorescent labeling of PCR fragments. Nat. Biotechnol. 18, 233-234.

Sherr, E.B., Sherr, B.F., Caron, D.A., Vaulot, D., Worden, A.Z., 2007. Oceanic Protists. Oceanography 20, 130-134.

578 Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M.D.M., Breiner, H.-W., Richards, T.A.,
579 2010. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex
580 eukaryotic community in marine anoxic water. Mol. Ecol. 19, 21-31.

581 Suikkanen, S., Laamanen, M., Huttunen, M., 2007. Long-term changes in summer
582 phytoplankton communities of the open northern Baltic Sea. Estuar. Coast. Shelf Sci. 71, 580-
583 592.

584 Telesh, I., Schubert, H., Skarlato, S., 2013. Life in the salinity gradient: Discovering
585 mechanisms behind a new biodiversity pattern. Estuarine Coastal and Shelf Science 135, 317-
586 327.

587 Telesh, I.V., Schubert, H., Skarlato, S.O., 2011. Revisiting Remane's concept: evidence for
588 high plankton diversity and a protistan species maximum in the horohalinicum of the Baltic
589 Sea. Mar. Ecol. Prog. Ser. 421, 1-11.

590 Telesh, I.V., Schubert, H., Skarlato, S.O., 2015. Size, seasonality, or salinity: What drives the
591 protistan species maximum in the horohalinicum? Estuar. Coast. Shelf Sci. 161, 102-111.

592 Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R., 2007. Naïve Bayesian Classifier for Rapid
593 Assignment of rRNA Sequences into the New Bacterial Taxonomy. Appl. Environ.
594 Microbiol. 73, 5261-5267.

595 Wasmund, N., Andrushaitis, A., Lysiak-Pastuszak, E., Muller-Karulis, B., Nausch, G.,
596 Neumann, T., Ojaveer, H., Olenina, I., Postel, L., Witek, Z., 2001. Trophic status of the south-

eastern Baltic Sea: A comparison of coastal and open areas. Estuarine Coastal and Shelf Science 53, 849-864.

Wasmund, N., Kownacka, J., Göbel, J., Jaanus, A., Johansen, M., Jurgensone, I., Lehtinen, S., Powilleit, M., 2017. The Diatom/Dinoflagellate Index as an Indicator of Ecosystem Changes in the Baltic Sea 1. Principle and Handling Instruction. Frontiers in Marine Science 4, 22.

Wielgat-Rychert, M., Ameryk, A., Jarosiewicz, A., Kownacka, J., Rychert, K., Szymanek, L., Zalewski, M., Agatova, A., Lapina, N., Torgunova, N., 2013. Impact of the inflow of Vistula river waters on the pelagic zone in the Gulf of Gdansk. Oceanologia 55, 859-886.

Witek, Z., Ochocki, S., Maciejowska, M., Pastuszak, M., Nakonieczny, J., Podgorska, B., Kownacka, J.M., Mackiewicz, T., Wrzesinska-Kwiecien, M., 1997a. Phytoplankton primary production and its utilization by the pelagic community in the coastal zone of the Gulf of Gdansk (southern Baltic). Mar. Ecol. Prog. Ser. 148, 169-186.

Witek, Z., Pastuszak, M., Nakonieczny, J., Ochocki, S., Zalewski, M., Maciejowska, M., Gromisz, S., Mackiewicz, T., Wrzesinska-Kwiecien, M., 1997b. Experimental studies on processes taking place in enclosures with mixed sea and Vistula River water. Oceanological Studies 1, 35-57.

Figures legends

Figure 1. A: Map of the Baltic Sea with the Gulf of Gdańsk marked by the rectangle. B: Location of the sampling stations at the Vistula River (F: freshwater site) and in the Gulf of Gdańsk (MZ: mixing zone site, B: brackish site). The position of the freshwater sampling station was fixed, while at sites MZ (S ~3.5) and B (S=7) sampling stations were selected based on the measured salinity.

Figure 2. Principal component analysis grouping the samples based on the environmental variables (triplicates showed for each site and date). The first principal component correlated positively with salinity and temperature, and negatively with concentrations of dissolved silica, total phosphorus and total nitrogen. The second principal component correlated positively with chlorophyll-*a* < 20 µm, and abundance of heterotrophic pico- and nanoplankton, and negatively with salinity.

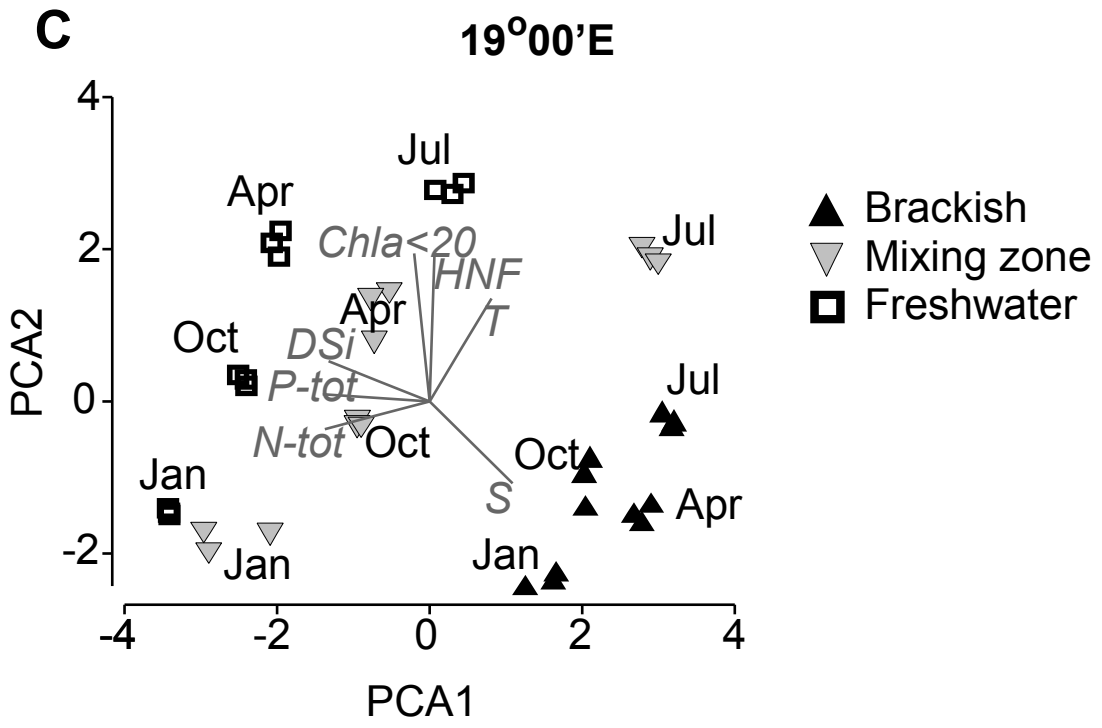
Figure 3. A: Species accumulation curve in the samples (lower X-axis), and rarefaction curve for all samples combined (upper X-axis); B: Numbers of observed phylotypes at the sampling sites in different months; C: Values of Shannon diversity index at the sampling sites in different months; D: Values of Pielou's evenness index at the sampling sites in different months. Average values± standard deviation (error bars) from triplicate samples are shown.

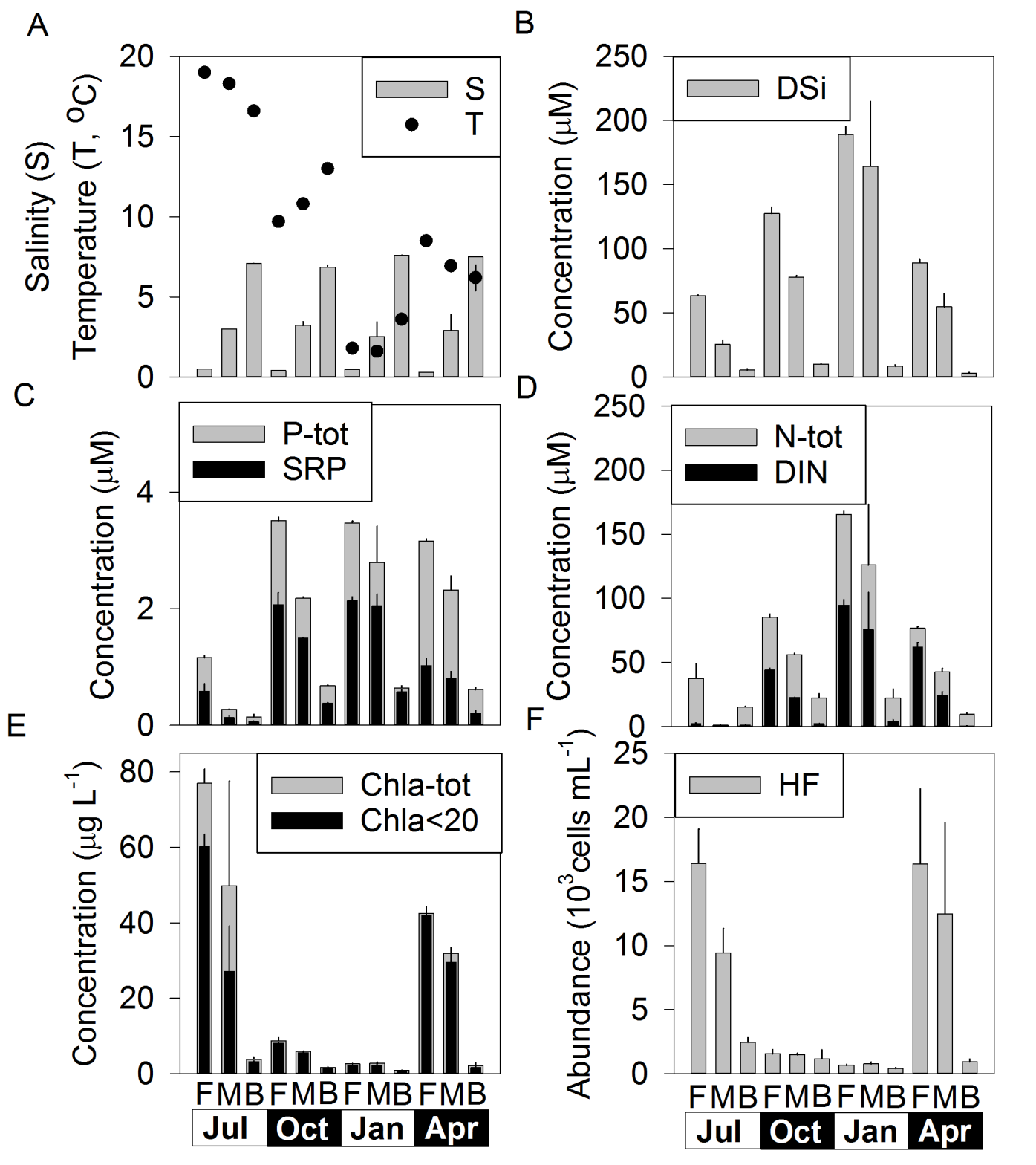
Figure 4. A: Venn diagrams showing number of unique and shared OTUs for all zones of the estuary over the sampling period; B: Fraction of reads coming from the main groups in different zones of the Vistula River estuary in different seasons. F – freshwater site, MZ – site in the mixing zone.

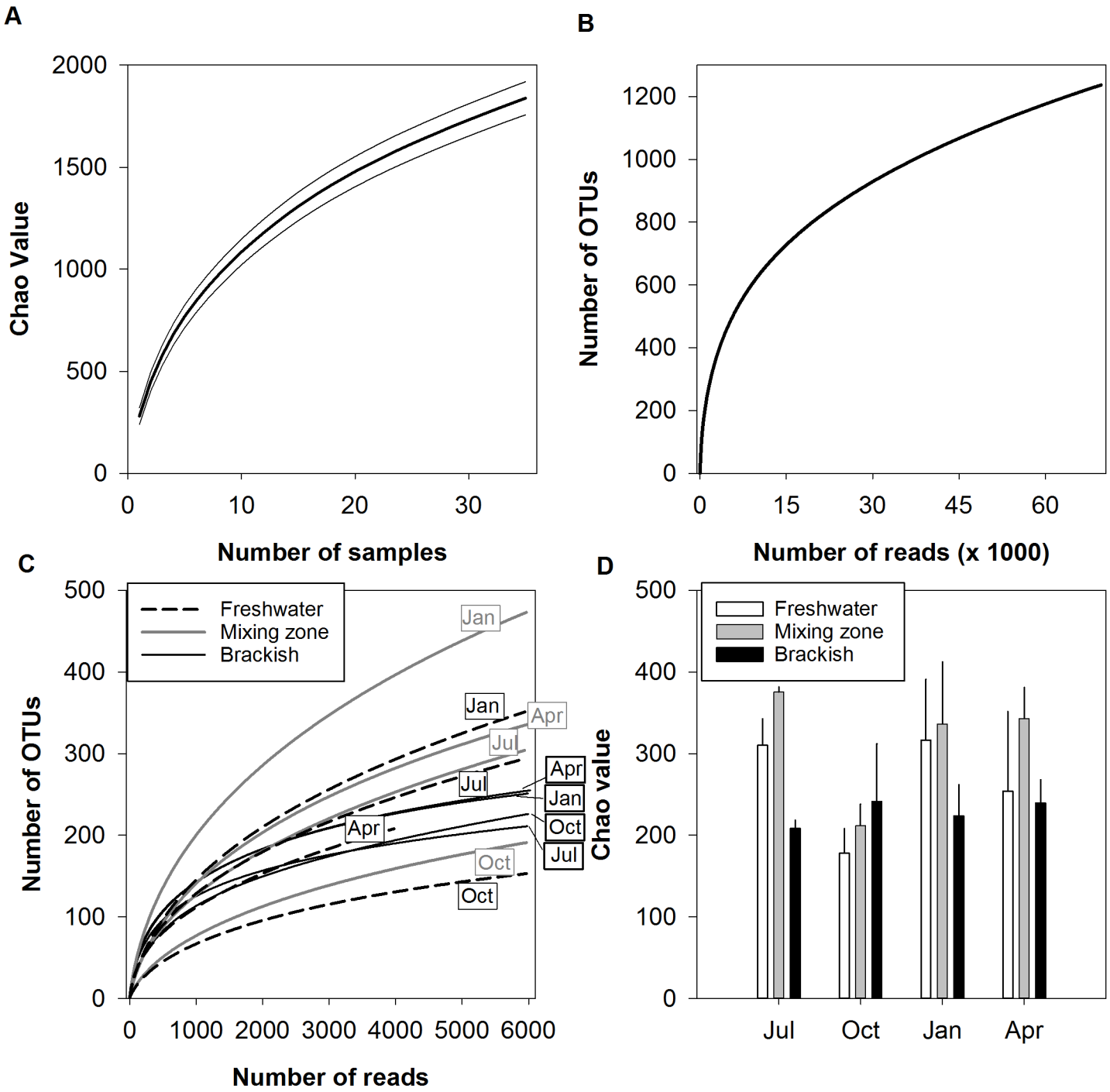
Figure 5. Ordination plot of distance-based redundancy analysis (dbRDA) relating the observed variability of pico- and nanoplanktonic communities to A: main taxonomic groups (lines). Only groups with Pearson coefficient > 0.7 are shown. 1 – Mamiellophyceae, 2 – Telonemia, 3 – Picobiliphyta, 4 – unclassified Cercozoa, 5 – Dictyochophyceae, 6 –

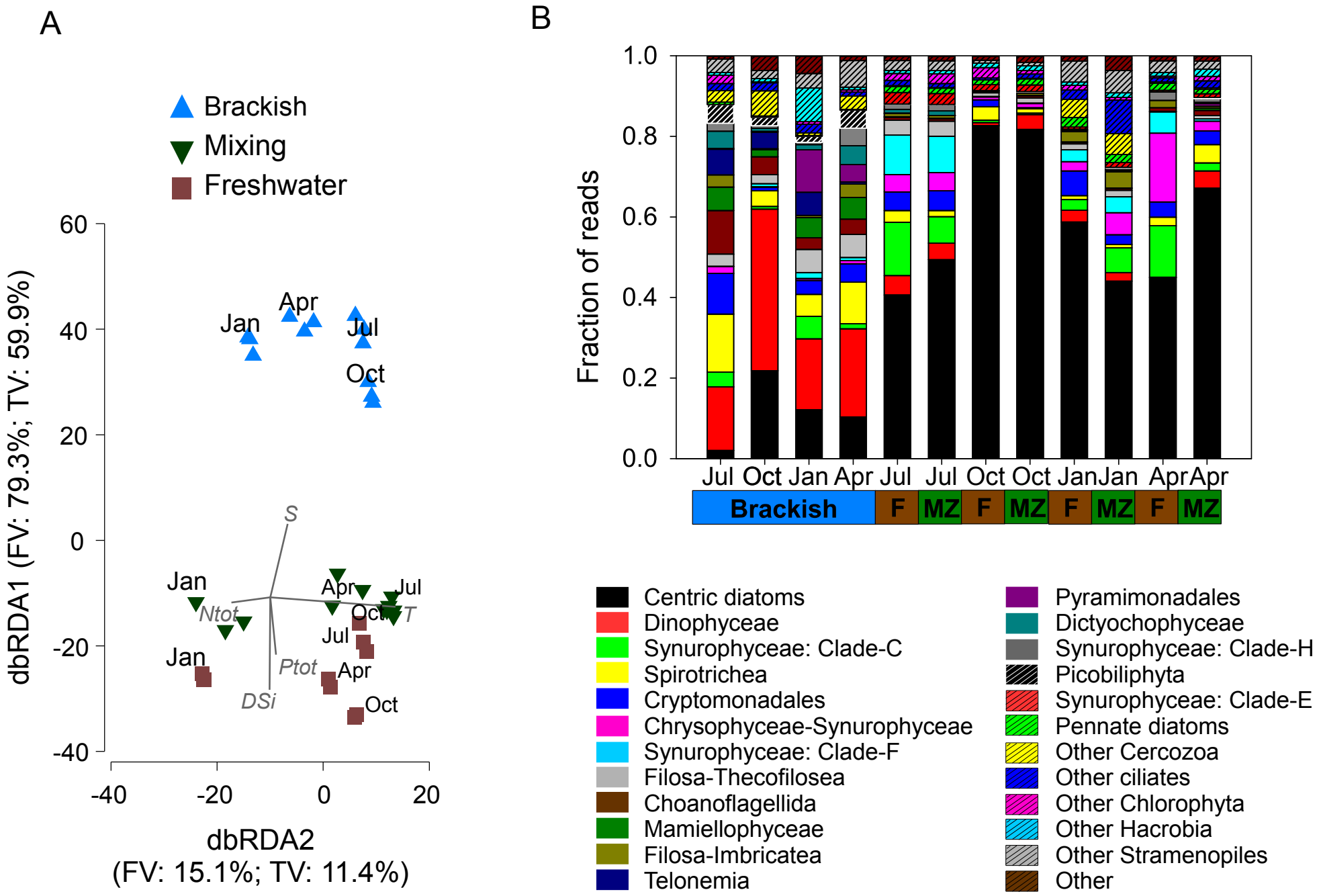
Dinophyceae, 7 – Spirotrichea, 8 – Mediophyceae, 9 – Raphid pennate, 10 – unclassified Opisthokonta, 11 – Coscinodiscophyceae. B: environmental explanatory variables (lines). Only statistically significant variables are shown. The proportions of the fitted variability (FV) and the total variability (TV) explained by the first two axes are given.

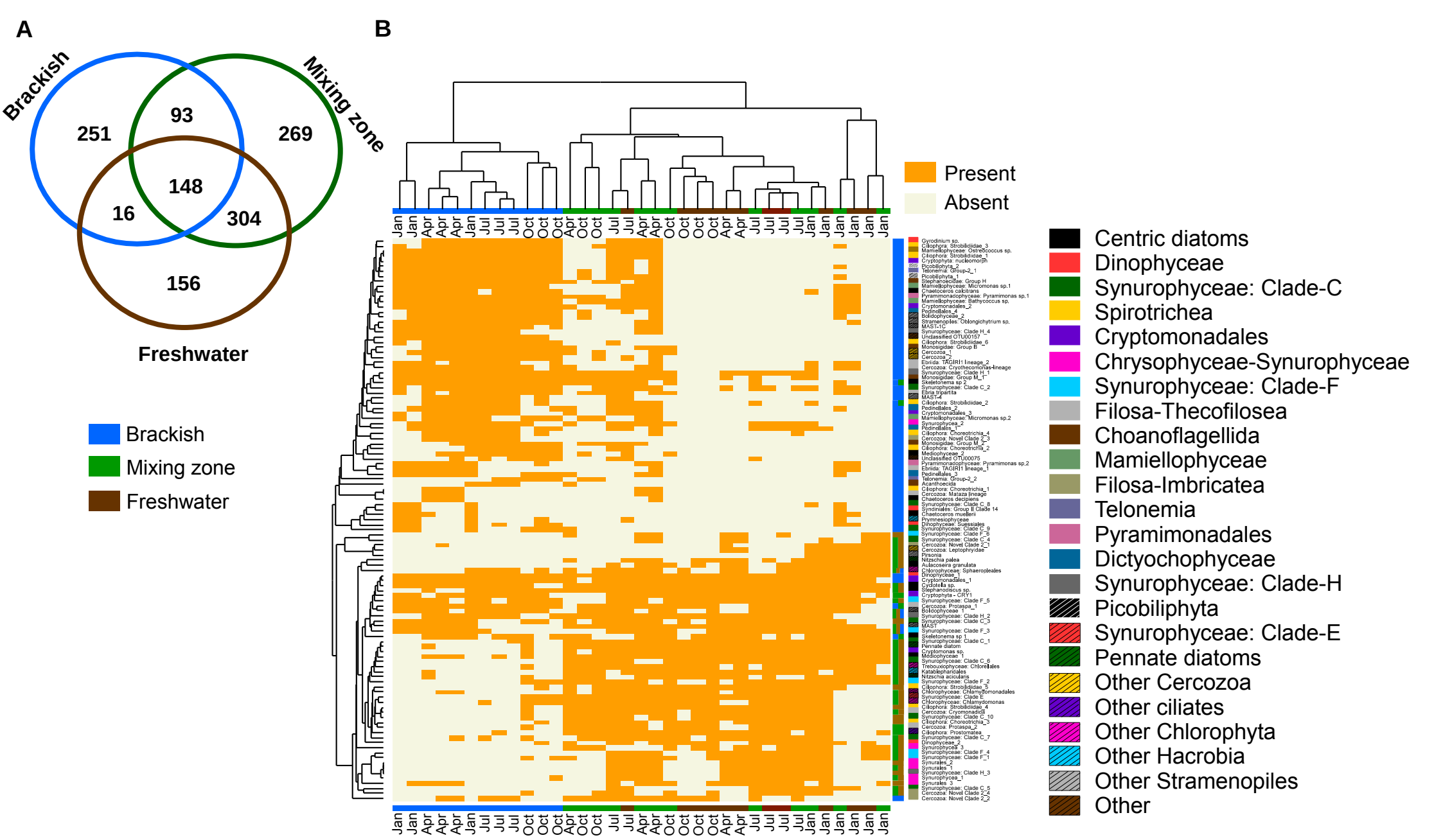

B
Gulf of Gdansk
A
Finland
Baltic Sea
Sweden
Estonia
Latvia
Lithuania
Denmark
Russia
Germany
Poland
Reda
B
54°30'N
MZ
Vistula Lagoon
Vistula
F


C
19°00'E
PCA2
PCA1
Jul
Apr
Chla<20
HNF
T
DSi
P-tot
N-tot
S
Oct
Jan
Brackish
Mixing zone
Freshwater


A
Salinity (S)
Temperature (T, °C)
S
T
B
Concentration (µM)
DSi
C
Concentration (µM)
P-tot
SRP
D
Concentration (µM)
N-tot
DIN
E
Concentration (µg L-1)
Chla-tot
Chla<20
F
Abundance (10^3cells mL^-1)
HF
F M B F M B F M B F M B
Jul
Oct
Jan
Apr


A
Chao Value
2000
1500
1000
500
0
0
10
20
30
Number of samples
B
Number of OTUs
1200
1000
800
600
400
200
0
0
15
30
45
60
Number of reads (x 1000)
C
Number of OTUs
500
400
300
200
100
0
Freshwater
Mixing zone
Brackish
Jan
Jan
Apr
Jul
Apr
Jul
Jan
Apr
Oct
Jul
Oct
Oct
0
1000
2000
3000
4000
5000
6000
Number of reads
D
Chao value
500
400
300
200
100
0
Freshwater
Mixing zone
Brackish
Jul
Oct
Jan
Apr


A
Brackish
Mixing
Freshwater
dbRDA1 (FV: 79.3%; TV: 59.9%)
dbRDA2
(FV: 15.1%; TV: 11.4%)
Jan
Apr
Jul
Oct
S
Ntot
T
Ptot
DSi
B
Fraction of reads
Brackish
F
MZ
Centric diatoms
Dinophyceae
Synurophyceae: Clade-C
Spirotrichea
Cryptomonadales
Chrysophyceae-Synurophyceae
Synurophyceae: Clade-F
Filosa-Thecofilosea
Choanoflagellida
Mamiellophyceae
Filosa-Imbricatea
Telonemia
Pyramimonadales
Dictyochophyceae
Synurophyceae: Clade-H
Picobiliphyta
Synurophyceae: Clade-E
Pennate diatoms
Other Cercozoa
Other ciliates
Other Chlorophyta
Other Hacrobia
Other Stramenopiles
Other


A
Brackish
Mixing zone
Freshwater
251
93
269
148
16
304
156
Brackish
Mixing zone
Freshwater
B
Present
Absent
Centric diatoms
Dinophyceae
Synurophyceae: Clade-C
Spirotrichea
Cryptomonadales
Chrysophyceae-Synurophyceae
Synurophyceae: Clade-F
Filosa-Thecofilosea
Choanoflagellida
Mamiellophyceae
Filosa-Imbricatea
Telonemia
Pyramimonadales
Dictyochophyceae
Synurophyceae: Clade-H
Picobiliphyta
Synurophyceae: Clade-E
Pennate diatoms
Other Cercozoa
Other ciliates
Other Chlorophyta
Other Hacrobia
Other Stramenopiles
Other

Kasia Piwosz, Joanna Całkiewicz, Marcin Gołębiewski, Simon Creer

Step by step: estuaries in brackish seas as possible zones of adaptation to different salinity regimes for pico- and nanoplanktonic protists.

Supplementary File 1: Methodology of 454 reads processing

In our study we applied high-throughput sequencing method (HTS) to detect phylotypes that were abundant in freshwater, and still present but rare at the brackish site, and thus to describe the microbiome of the whole estuary in more detail. We utilised 454 sequencing of the V3-V4 rRNA fragment, whose length (500-650 bp) facilitated opportunities for more detailed phylogenetic analysis and the detection of subOTUs occuring in different habitats. To mitigate the possible problems arising from errors during demultiplexing the reads (wrong assignment of reads to samples), we used a set of long barcodes (10 nt) with minimal edit distance equal to 4 and did allow only one mismatch in a barcode. As the error probability in raw reads is close to 1e-3, the probability of erroneous read assignment due to one tag mutating into another is 1e-09 under assumption of independent mutations. Thus, chimera formation might be the only mechanism leading to tag misidentification in our case and, as we employed three-step chimera removal procedure, it might be safely assumed that the number of misidentified tags in our data was negligible.

The flows were extracted from the .sff files, forward and reverse reads separately (sffinfo), then they were assigned to the samples basing on the MID sequences, trimmed to min. 500 and max. 650 flows (trim.flows) and denoised with AmpliconNoise algorithms (shhh.flows and shhh.seqs). Primers and MIDs were removed from the denoised seuqences, the reverse reads were reverse complemented (trim.seqs), and the reads set was dereplicated (unique.seqs). The forward and reverse read sets were pooled (cat) and the whole set was dereplicated again and aligned to the

SILVA template alignment (align.seqs). Reads covering the desired region of the alignment (pos. 13900-22400) were chosen (screen.seqs) and gap only and terminal gap-containing columns were removed from the alignment (filter.seqs). The set was dereplicated again and residual sequencing and PCR noise was removed with Single Linkage pre-clustering (pre.cluster, Huse *et al*., 2010). Chimera identification and removal was performed in three rounds: i) with UCHIME (chimera.uchime, Edgar *et al*., 2011), ii) with PERSEUS (chimera.perseus, Quince *et al*., 2011) and iii) with chimera slayer (chimera.slayer, Haas et al., 2011) using the PR2 template alignment prepared by aligning the sequences to the SILVA template and screening for sequences covering the same region of the alignment as the reads (13900-22400) .

To increase taxonomic resolution, full-length sequences (list.seqs, get.seqs) were used for classification with a naive Bayesian classifier (classify.seqs, Wang *et al*., 2007) with the PR2 template and taxonomy files (downloaded at http://ssu-rrna.org/pr2 on May 14, 2014) at the bootstrap confidence level of 80%. Taxa assigned as 'unknown' were removed from the final set. Average linkage (UPGMA) algorithm was used to construct OTUs at the 0.03 dissimilarity level, and singletons as well as doubletons were removed from the data (remove.rare).

To ensure that OTUs frequencies in the subsampled dataset are close to the original ones, the final reads set was subsampled ten times to 2500 reads per sample (sub.sample), read names were mangled to reflect their coming from a particular subsample (regular expressions in the sed editor), subsamples were combined (cat), the whole set was dereplicated and used for distance matrix calculation (dist.seqs) and OTU construction via average neighbor clustering at 97% similarity level (cluster). A shared OTU table was constructed (make.shared) and the table averaged over the subsamples (i.e. for each OTU numbers of reads found in each subsample were summed and the sum was divided by ten) was calculated with a Perl script (average_shared.perl). OTUs were classified using consensus approach with PR2 taxonomic assignment (classify.otu).

Details are given below:

#Prerequisites: Mothur 1.32 installed under Linux environment (executable present in a directory listed in $PATH is assumed) , Lookup_Titanium.pat in a directory visible for mothur, SILVA and PR2 files in a directory visible for mothur, bash shell, vi and sed editors, Perl 5, sff files, oligos files with samples assignment.

#Lines starting with # are commentaries, other lines are code to be copied to a terminal.

x, x1, etc. denote a generic filename.

#In mothur commands the number of processors can (and should) be changed to be lower than the number of accessible processors

#cd to the directory where sff files are stored

```
mkdir forward reverse
```

```
mothur
```

#For each sff file execute:

```
sff.info(sff=x.sff, flow=T)
quit()
```

```
cd forward
```

#For each flow file execute:

```
ln -s ../x.flow .
```

#Start mothur:

```
mothur
```

#For each flow file execute:

trim.flows(flow=x.flow, oligos=x_f.oligos, pdiffs=2, bdiffs=1, processors=6)

shhh.flows(file=x.flow.files, processors=18)

shhh.seqs(fasta=x.shhh.fasta, name=x.shhh.names, group=x.shhh.groups)

#Include files derived from all sffs

trim.seqs(fasta=x.shhh.shhh_seqs.fasta, name=x.shhh.shhh_seqs.names, oligos=x_f.oligos,

pdiffs=2, bdiffs=1, processors=4)

system(cat x.shhh.shhh_seqs.trim.fasta x1.shhh.shhh_seqs.trim.fasta x2.shhh.shhh_seqs.trim.fasta >

eukarya_f.shhh.shhh_seqs.trim.fasta)

system(cat x.shhh.shhh_seqs.trim.names x1.shhh.shhh_seqs.trim.names

x2.shhh.shhh_seqs.trim.names > eukarya_f.shhh.shhh_seqs.trim.names)

system(cat x.shhh.shhh_seqs.groups x1.shhh.shhh_seqs.groups x2.shhh.shhh_seqs.groups >

eukarya_f.shhh.shhh_seqs.groups)

quit()

cd ../reverse

#For each flow file execute:

ln -s ../x.flow .

#Start mothur:

mothur

#For each flow file execute:

trim.flows(flow=x.flow, oligos=x_r.oligos, pdiffs=2, bdiffs=1, processors=6)

shhh.flows(file=x.flow.files, processors=18)

shhh.seqs(fasta=x.shhh.fasta, name=x.shhh.names, group=x.shhh.groups, processors=1)

#Include files derived from all sffs

trim.seqs(fasta=x.shhh.shhh_seqs.fasta, name=x.shhh.shhh_seqs.names, oligos=x_f.oligos,

pdiffs=2, bdiffs=1, reverse=T, processors=4)

system(cat x.shhh.shhh_seqs.trim.fasta x1.shhh.shhh_seqs.trim.fasta x2.shhh.shhh_seqs.trim.fasta >

eukarya_r.shhh.shhh_seqs.trim.fasta)

system(cat x.shhh.shhh_seqs.trim.names x1.shhh.shhh_seqs.trim.names

x2.shhh.shhh_seqs.trim.names > eukarya_r.shhh.shhh_seqs.trim.names)

system(cat x.shhh.shhh_seqs.groups x1.shhh.shhh_seqs.groups x2.shhh.shhh_seqs.groups >

eukarya_r.shhh.shhh_seqs.groups)

quit()

cd ..

cat forward/eukarya_f.shhh.shhh_seqs.fasta reverse/eukarya_r.shhh.shhh_seqs.fasta > eukarya.fasta

cat forward/eukarya_f.shhh.shhh_seqs.names reverse/eukarya_r.shhh.shhh_seqs.names >

eukarya.names

cat forward/eukarya_f.shhh.shhh_seqs.groups reverse/eukarya_f.shhh.shhh_seqs.groups >

eukarya.groups

mothur

unique.seqs(fasta=eukarya.fasta, name=eukarya.names)

align.seqs(fasta=current, reference=silva.eukarya.fasta, processors=16)

remove.seqs(fasta=current, name=current, group=eukarya.groups, accnos=current)

screen.seqs(fasta=current, name=current, group=current, start=6500, end=22500)

filter.seqs(fasta=current, vertical=T, trump=.)

unique.seqs(fasta=current, name=current)

pre.cluster(fasta=current, name=current, group=current)

chimera.uchime(fasta=current, name=current, group=current, reference=groups)

remove.seqs(fasta=current, name=current, group=current, accnos=current)

chimera.perseus(fasta=current, name=current, group=current)

remove.seqs(fasta=current, name=current, group=current, accnos=current)

chimera.slayer(fasta=current, name=current, group=current, reference=pr2.good.filter.pick.ng.fasta)

remove.seqs(fasta=current, name=current, group=current, accnos=current)

list.seqs(fasta=current)

get.seqs(fasta=eukarya.fasta, accnos=current) #get full length seqs for classification

classify.seqs(fasta=current, name=current, group=current, reference=pr2.good.filter.pick.ng.fasta,

taxonomy=pr2.pick.tax, cutoff=80)

remove.lineage(fasta=current, name=current, group=current, taxonomy=current, taxon=unknown;)

dist.seqs(fasta=current, cutoff=0.10, processors=16)

cluster(column=current, name=current)

remove.rare(list=current, label=0.03, nseqs=2)

list.seqs(list=current)

get.seqs(fasta=current, name=current, group=current, accnos=current) #get seqs set without

singletons and doubletons

quit()

mv eukarya.unique.pick.good.filter.unique.precluster.pick.pick.pick.pick.pick.fasta

eukarya.final.fasta

mv eukarya.unique.pick.good.filter.unique.precluster.pick.pick.pick.pick.pick.names eukarya.final.names

mv eukarya.pick.good.pick.pick.pick.pick.pick.groups eukarya.final.groups

mv eukarya.unique.pick.good.filter.unique.precluster.pick.pick.pick.pick.pick.0.03.an.pick.list eukarya.final.an.list

#The procedure below was devised to mitigate the effect of single subsampling, namely possibility of OTU frequencies being far off the real ones (meaning the frequencies in the whole dataset). Ten subsamples are generated, read names are mangled to reflect their coming from a particular subsample, the resulting set is dereplicated and OTUs are constructed as above. Shared OTU table is then constructed and averaged over the subsamples (i.e. numbers of reads coming from a given OTU in each subsample are summed and the result is divided by the number of subsamples). The reads are classified and the results are averaged analogically, but at taxa levels instead of OTUs. There is a possibility of bootstrapping in some mothur commands, such as unifrac.(un)weighted, summary.single or dist.shared. Its was used here.

for f in 1 2 3 4 5 6 7 8 9 10; do mothur „#sub.sample(fasta=eukarya.final.fasta, name=eukarya.final.names, group=eukarya.final.groups, pergroup=T, size=2500);”; cat eukarya.final.subsample.fasta | sed „s/>/>$f_/” >> eukarya.bootstrap.fasta; cat eukarya.final.subsample.names | sed „s/^/$_/” | sed „s/\t/\t$f_/” | sed „s/\,/\,$f_/g” >> eukarya.bootstrap.names; cat eukarya.final.subsample.groups | sed „s/^/$f\_/” | sed „s/$/_$f/” >> eukarya.bootstrap.groups; done

mothur

unique.seqs(fasta=eukarya.bootstrap.fasta, name=eukarya.bootstrap.names)

list.seqs(fasta=eukarya.bootstrap.unique.fasta)

dist.seqs(fasta=current, cutoff=0.10, processors=16)

cluster(column=current, name=current)

make.shared(list=current, group=eukarya.bootstrap.groups, label=0.03) #shared OTU table for averaging

make.shared(list=eukarya.final.an.list, group=eukarya.final.groups, label=0.03) #shared OTU table for diversity estimations and generation of community distance matrices

dist.shared(shared=current, calc=braycurtis-morisitahorn, subsample=2500, iters=100)

summary.single(shared=current, calc=sobs-chao-ace-shannon-shannoneven, subsample=2500, iters=100)

clearcut(fasta=eukarya.final.fasta, DNA=T, kimura=T)

unifrac.weighted(tree=current, name=eukarya.final.names, group=eukarya.final.groups, subsample=2500, distance=lt, processors=16)

quit()

extract_full_length_seqs.perl -l eukarya.bootstrap.unique.accnos -f eukarya.fasta > eukarya.bootstrap.unique.fullength.fasta #the script fetches sequences from a fasta file whose names are those from the accnos file with subsample number dropped, sequences from the fasta file are printed with names coming from the accnos file

mothur

classify.seqs(fasta=eukarya.bootstrap.unique.fullength.fasta, name=eukarya.bootstrap.unique.names, group=eukarya.bootstrap.groups, reference=pr2.good.filter.pick.ng.fasta, taxonomy=pr2.pick.tax, cutoff=80, probs=F, processors=16) #no bootstrap probabilities, they preclude OTUs classification with classify.otu

```
classify.otu(list=eukarya.bootstrap.unique.an.list,
taxonomy=eukarya.bootstrap.unique.fullength.wang.taxonomy,
name=eukarya.bootstrap.unique.names, cutoff=80)
quit()

average_shared.perl eukarya.bootstrap.unique.an.shared > eukarya.bootstrap.unique.an.avg.shared

average_tax.summary.perl -f eukarya.bootstrap.unique.fullength.wang.tax.summary -n 10 >
eukarya.bootstrap.unique.fullength.avg.tax.summary.csv

#For vegan-based analyses the shared OTUs file was manually edited in vi to remove a redundant
tabulator at the end of the header line and was imported to R

R
eukarya.community <- read.table(„eukarya.bootstrap.unique.an.avg.shared”, header=T, sep=”\t”,
dec=”.”)
rownames(eukarya.community) <- eukarya.community$Group
eukarya.community$Group <- NULL
eukarya.community$label <- NULL
eukarya.community$numOtus <- NULL

#Construction for subOTUs for 50 most abundant OTUs
for f in {1..50}; do  get_otu_reads_accnos.perl eukarya.bootstrap.unique.an.list 0.03 $f >
otu$f\.accnos; mothur „#get.seqs(fasta=eukarya.bootstrap.unique.fasta,
name=eukarya.bootstrap.unique.names, group=eukarya.bootstrap.groups, accnos=otu$f\.accnos);”;
```

mv eukarya.bootstrap.unique.pick.fasta otu$f\.fasta; mv eukarya.bootstrap.unique.pick.names otu$f\.names; mv eukarya.bootstrap.pick.groups otu$f\.groups; mothur „#dist.seqs(fasta=otu$f\.fasta, cutoff=0.10, processors=4); cluster(column=otu$f\.dist, name=otu$f\.names); make.shared(list=otu$f\.an.list, group=otu$f\.groups, label=0.01); get.oturep(list=otu$f\.an.list, column=otu$f\.dist, name=otu$f\.names, fasta=otu$f\.fasta, label=0.01, method=distance, weighted=T); clearcut(fasta=otu$f\.an.0.01.rep.fasta, DNA=T, kimura=T);"; cat otu$f\.an.0.01.rep.tre | sed „s/Otu/subOtu/g" > otu$f\.an.0.01.rep.mod.tre; cat otu$f.an.shared | sed „s/Otu/subOtu/g" > otu$f\.an.mod.shared; done

#Trees generated by the version of clearcut incorporated into mothur are sometimes not conforming to the standard and need to be manually edited to be correctly read by phyloseq's import_mothur function. The problem lies in an unnecessary pair of parentheses, where the closing one directly precedes a comma. This pair should be removed.
#Sample data file should be prepared as a tab-separated file. The file should include site and season for each sample.

R

library(phyloseq)

sdata ← read.table(„sample_data.csv", header=T, sep="\t");

sdata$site ← factor(sdata$site, levels=c('freshwater','mixing_zone', 'brackish'))

sdata$season ← factor(sdata$season, levels=c('spring', 'summer', 'autumn', 'winter'))

#For each OTU execute

otux ← import_mothur(mothur_shared_file="otux.an.mod.shared",

mothur_tree_file="otux.an.0.01.rep.tre", cutoff=0.01)

sample_data(otux) ← sample_data(sdata)

pdf(file="otux.pdf")

```
print( plot_tree(otux, shape=”season”, color=”site”, size=”abundance”, label.tips=”taxa_names”,
title=”OTUx”) )
dev.off()

#The pdf files may be collated later, or printing may be performed within a 'for' loop with
pdf(file=”...”, onefile=T)
```

Supplementary Table S1. Fractions of freshwater and brackish water in the mixing zone, calculated based on salinity*, and deviations of the theoretical values of environmental variables from the measured values (in percent relative to the measured values), computed from fractions of freshwater in the mixing zone. Vistula waters: fraction of the freshwaters in the mixing zone;, Brackish waters: fraction of the brackish waters in the mixing zone; P-tot – total phosphorus, N-tot – total nitrogen, DSi – dissolved silica in µM, Chl-a –chlorophyll-*a*. For details see Golebiewski et al.

.

Date	**Vistula waters**	**Brackish waters**	**P_{tot}**	**N_{tot}**	**DSi**	**Chl-*a***
07 Jul 2011	0.62	0.38	-175.75	-2178.29	-62.73	1.28
19 Oct 2011	0.56	0.44	-3.75	-2.91	2.15	3.48
25 Jan 2012	0.70	0.30	5.58	2.53	17.62	25.22
17 Apr 2012	0.64	0.36	3.36	-22.98	-5.86	12.57

*The proportion of fresh waters was calculated as:

1. $fr = \frac{\mathrm{Sm}-\mathrm{S}b}{\mathrm{Sr}-\mathrm{Sb}}$

and of brackish waters as:

2. $fb = 1 - fr$,

where:

fr – fraction of freshwater;

fb – fraction of brackish water;

Sm – salinity in the mixing zone;

Sb – salinity at the brackish site;

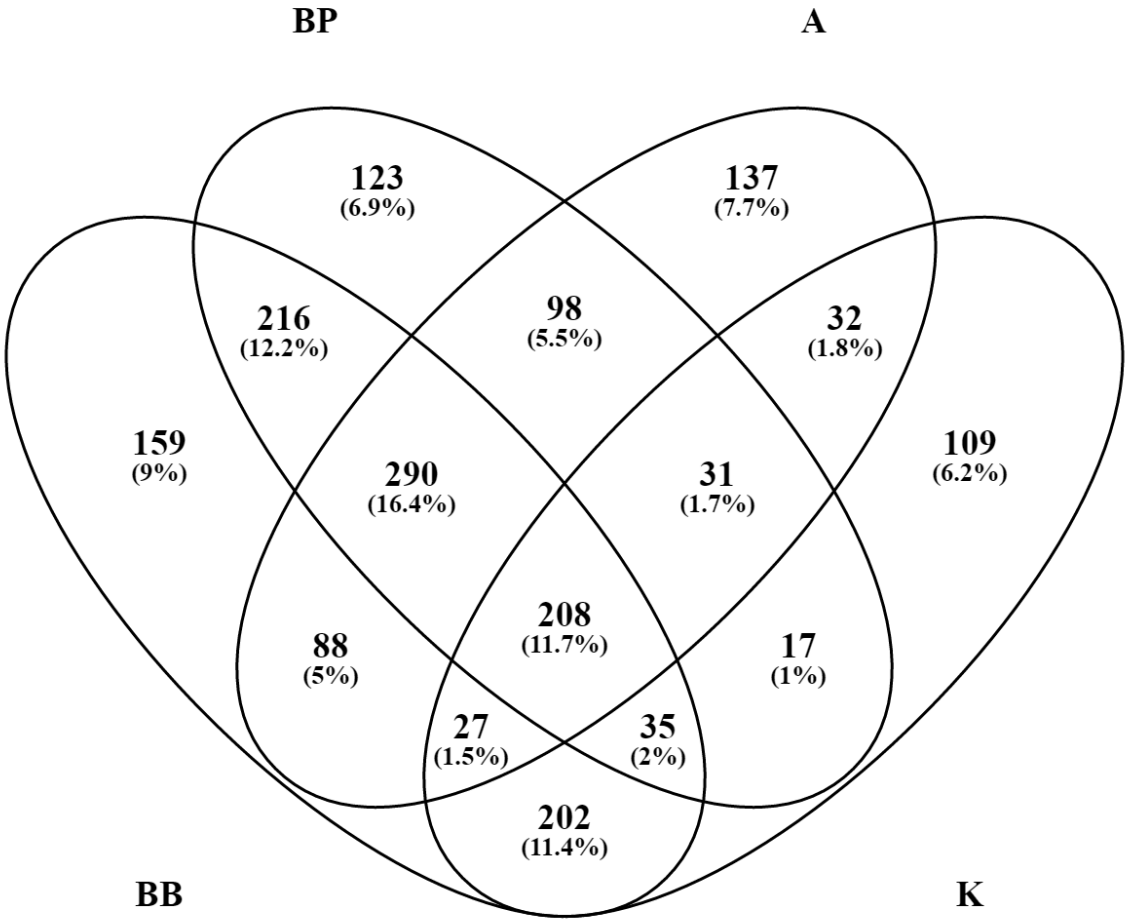
Sr – salinity at the freshwater site.

Evaporation and precipitation were assumed negligible (Ameryk *et al*., 2005).

Supplementary Table S3: Classification of reads according to the site of their peak abundance (type). Summary number of reads for each zone and for each zone and season, toghether with taxonomic affiliations, are also given.
Core OTUs at treshhold 1% are marked in bold and italics; core OTUs at treshhold 0.5% are marked in italics

	summary abundance for all seasons				Brackish				Mixing Zone				Freshwater											
ID	Brackish	Mixing Zone	Freshwater	Habitat	summer (Jul'11)	autumn (Oct'11)	winter (Jan'12)	spring (Apr'12)	summer (Jul'11)	autumn (Oct'11)	winter (Jan'12)	spring (Apr'12)	summer (Jul'11)	autumn (Oct'11)	winter (Jan'12)	spring (Apr'12)	total sumof reads	taxonomy						
OTU4	***5040***	***637***	***376***	***Brackish***	***872***	***2253***	***924***	***991***	***165***	***216***	***76***	***180***	***216***	***38***	***119***	***3***	***6053***	***Alveolata(100)***	***Dinophyta(100)***	***Dinophyceae(100)***	***Dinophyceae_X(100)***	***Dinophyceae_XX(92)***	***unclassified(85)***	***unclassified(85)***
OTU12	851	4	5	Brackish	347	170	304	30	1	2	0	1	5	0	0	0	860	Hacrobia(100)	Telonemia(100)	Telonemia_X(100)	Telonemia_XX(100)	Telonemia-Group-2(100)	Telonemia-Group-2_X(100)	Telonemia-Group-2_X+sp.(100)
OTU14	***574***	***139***	***92***	***Brackish***	***267***	***27***	***120***	***160***	***71***	***3***	***15***	***50***	***56***	***17***	***17***	***2***	***805***	***Hacrobia(100)***	***Cryptophyta(100)***	***Cryptophyceae(100)***	***Cryptophyceae_X(100)***	***Cryptomonadales(100)***	***Teleaulax(87)***	***Teleaulax+amphioxeia(87)***
OTU22	532	30	4	Brackish	13	10	298	211	0	2	1	27	1	0	3	0	566	Archaeplastida(100)	Chlorophyta(100)	Pyramimonadales(100)	Pyramimonadales_X(100)	Pyramimonadales_XX(100)	Pyramimonas(100)	unclassified(100)
OTU23	460	25	8	Brackish	192	46	112	110	5	3	1	16	7	0	1	0	493	Archaeplastida(100)	Chlorophyta(100)	Mamiellophyceae(100)	Mamiellales(100)	Bathycoccaceae(100)	Bathycoccus(100)	Bathycoccus+prasinos(96)
OTU25	449	4	3	Brackish	266	16	157	10	1	0	0	3	3	0	0	0	456	Alveolata(100)	Ciliophora(100)	Spirotrichea(100)	Choreotrichia(100)	Strobilidiidae(100)	Strobilidiidae_X(100)	Strobilidiidae_X+sp.(100)
OTU27	430	6	7	Brackish	4	11	388	27	0	0	3	3	1	0	6	0	443	Stramenopiles(100)	Stramenopiles_X(100)	Bacillariophyta(100)	Bacillariophyta_X(100)	Polar-centric-Mediophyceae(100)	Chaetoceros(100)	Chaetoceros+calcitrans(100)
OTU30	398	18	4	Brackish	127	78	48	145	2	3	1	12	4	0	0	0	420	Hacrobia(100)	Picobiliphyta(100)	Picobiliphyta_X(100)	Picobiliphyta_XX(100)	Picobiliphyta_XXX(100)	Picobiliphyta_XXXX(100)	Picobiliphyta_XXXX+sp.(100)
OTU33	344	72	1	Brackish	20	55	6	263	1	0	0	71	1	0	0	0	417	Alveolata(100)	Dinophyta(100)	Dinophyceae(100)	Dinophyceae_X(100)	Dinophyceae_XX(100)	Gyrodinium(100)	Gyrodinium+dominans(81)
OTU37	361	24	8	Brackish	231	58	57	15	5	9	3	7	7	0	0	1	393	Opisthokonta(100)	Choanoflagellida(100)	Choanoflagellatea(100)	Craspedida(100)	Monosigidae_Group_M(100)	Monosigidae_Group_M_X(100)	Monosigidae_Group_M_X+sp.(100)
OTU39	324	24	13	Brackish	0	2	292	30	0	0	14	10	0	0	13	0	361	Archaeplastida(100)	Chlorophyta(100)	Pyramimonadales(100)	Pyramimonadales_X(100)	Pyramimonadales_XX(100)	Pyramimonas(100)	Pyramimonas+sp.(100)
OTU41	300	27	14	Brackish	165	18	64	53	12	0	2	13	12	0	2	0	341	Hacrobia(100)	Cryptophyta(100)	Cryptophyceae(100)	Cryptophyceae_X(100)	Cryptomonadales(100)	unclassified	unclassified
OTU43	323	2	3	Brackish	163	137	23	0	2	0	0	0	3	0	0	0	328	Alveolata(100)	Ciliophora(100)	Spirotrichea(100)	Choreotrichia(100)	Choreotrichia-1(100)	Choreotrichia-1_X(100)	Choreotrichia-1_X+sp.(100)
OTU47	288	10	2	Brackish	101	47	40	100	2	0	1	7	2	0	0	0	300	Hacrobia(100)	Picobiliphyta(100)	Picobiliphyta_X(100)	Picobiliphyta_XX(100)	Picobiliphyta_XXX(100)	Picobiliphyta_XXXX(100)	Picobiliphyta_XXXX+sp.(100)
OTU48	283	10	5	Brackish	64	3	121	95	0	0	3	7	1	0	4	0	298	Archaeplastida(100)	Chlorophyta(100)	Mamiellophyceae(100)	Mamiellales(100)	Mamiellaceae(100)	Micromonas(100)	unclassified(100)
OTU52	284	5	1	Brackish	0	1	283	0	0	0	5	0	1	0	0	0	290	Hacrobia(100)	Haptophyta(100)	Prymnesiophyceae(100)	Prymnesiales(100)	Prymnesiaceae(100)	unclassified	unclassified
OTU56	270	8	4	Brackish	191	42	36	1	1	5	0	2	4	0	0	0	282	Opisthokonta(100)	Choanoflagellida(100)	Choanoflagellatea(100)	Craspedida(100)	Monosigidae_Group_M(100)	Monosigidae_Group_M_X(100)	Monosigidae_Group_M_X+sp.(100)
OTU57	260	13	7	Brackish	198	2	20	40	3	0	1	9	7	0	0	0	280	Stramenopiles(100)	Stramenopiles_X(100)	Dictyochophyceae(100)	Dictyochophyceae_X(100)	Pedinellales(100)	Pedinellales_X(100)	Pedinellales_X+sp.(100)
OTU60	222	28	22	Brackish	47	27	19	129	5	3	1	19	4	0	0	18	272	Stramenopiles(100)	Stramenopiles_X(100)	Chrysophyceae-Synurophyceae(100)	Chrysophyceae-Synurophyceae_X(100)	Clade-H(100)	Clade-H_X(100)	Clade-H_X+sp.(100)
OTU63	218	23	7	Brackish	153	30	28	7	12	1	4	6	7	0	0	0	248	Alveolata(100)	Ciliophora(100)	Spirotrichea(100)	Oligotrichia(100)	Strombidiidae(100)	unclassified(83)	unclassified(83)
OTU64	231	7	6	Brackish	0	0	230	1	0	0	7	0	0	0	6	0	244	Stramenopiles(100)	Stramenopiles_X(100)	Bacillariophyta(100)	Bacillariophyta_X(100)	Polar-centric-Mediophyceae(100)	Chaetoceros(100)	Chaetoceros+muellerii(100)
OTU75	200	2	6	Brackish	98	101	0	1	2	0	0	0	6	0	0	0	208	unclassified(100)	unclassified(100)	unclassified(100)	unclassified(100)	unclassified(100)	unclassified(100)	unclassified(100)
OTU78	141	37	25	Brackish	135	0	0	6	8	26	0	3	12	13	0	0	203	Rhizaria(100)	Cercozoa(100)	Filosa-Imbricatea(100)	Filosa-Imbricatea_X(100)	Novel-clade-2(100)	Novel-clade-2_X(100)	Novel-clade-2_X+sp.(100)
OTU79	181	15	2	Brackish	124	0	16	41	4	0	0	11	2	0	0	0	198	Hacrobia(100)	Cryptophyta(100)	Cryptophyceae(100)	Cryptophyceae_X(100)	Cryptomonadales(100)	unclassified(93)	unclassified(93)
OTU90	*92*	*39*	*36*	*Brackish*	*11*	*4*	*61*	*16*	*0*	*0*	*34*	*5*	*0*	*0*	*36*	*0*	*167*	*Rhizaria(100)*	*Cercozoa(100)*	*Filosa-Thecofilosea(100)*	*Cryomonadida(100)*	*Cryothecomonas-lineage(95)*	*Cryothecomonas-lineage_X(95)*	*Cryothecomonas-lineage_X+sp.(95)*
OTU92	150	7	7	Brackish	68	5	53	24	2	0	0	5	7	0	0	0	164	Archaeplastida(100)	Chlorophyta(100)	Mamiellophyceae(100)	Mamiellales(100)	Bathycoccaceae(100)	Ostreococcus(100)	Ostreococcus+sp.(100)
OTU95	153	3	2	Brackish	31	80	3	39	0	1	0	2	2	0	0	0	158	Rhizaria(100)	Cercozoa(100)	Cercozoa_X(100)	Cercozoa_XX(100)	Cercozoa_XXX(100)	Cercozoa_XXXX(100)	Cercozoa_XXXX+sp.(100)
OTU97	146	7	1	Brackish	3	14	43	86	0	0	0	7	0	0	1	0	154	Stramenopiles(100)	Stramenopiles_X(100)	Stramenopiles_XX(100)	Stramenopiles_XXX(100)	Stramenopiles_XXXX(100)	Oblongichytrium(100)	Oblongichytrium+sp.(100)
OTU99	135	5	3	Brackish	0	0	124	11	0	0	2	3	0	0	3	0	143	Rhizaria(100)	Cercozoa(100)	Filosa-Thecofilosea(100)	Ebriida(100)	TAGIRI1-lineage(100)	TAGIRI1-lineage_X(100)	TAGIRI1-lineage_X+sp.(100)
OTU102	110	17	10	Brackish	73	1	8	28	9	1	2	5	10	0	0	0	137	Stramenopiles(100)	Stramenopiles_X(100)	Chrysophyceae-Synurophyceae(100)	Chrysophyceae-Synurophyceae_X(88)	unclassified(88)	unclassified(88)	unclassified(88)
OTU105	130	3	1	Brackish	80	12	26	12	2	0	0	1	1	0	0	0	134	nucleomorph-Archaeplastida(100)	Cryptophyta-nucleomorph(100)	Cryptophyta-nucleomorph_X(100)	Cryptophyta-nucleomorph_XX(100)	Cryptophyta-nucleomorph_XXX(100)	Cryptophyta-nucleomorph_XXXX(100)	Cryptophyta-nucleomorph_XXXX+sp.(100)
OTU107	105	17	4	Brackish	16	82	3	4	3	10	3	1	3	0	1	0	126	Rhizaria(100)	Cercozoa(100)	Filosa-Thecofilosea(100)	Ebriida(100)	Ebriidae(100)	Ebria(100)	Ebria+tripartita(100)
OTU108	114	2	3	Brackish	72	20	17	5	0	2	0	0	3	0	0	0	119	Stramenopiles(100)	Stramenopiles_X(100)	Bacillariophyta(100)	Bacillariophyta_X(100)	Polar-centric-Mediophyceae(100)	unclassified(100)	unclassified(100)
OTU121	104	1	1	Brackish	74	1	7	22	1	0	0	0	1	0	0	0	106	Stramenopiles(100)	Stramenopiles_X(100)	MAST(100)	MAST-4-6-7-8-9-10-11(100)	MAST-4(100)	MAST-4_X(100)	MAST-4_X+sp.(100)
OTU127	97	1	1	Brackish	1	2	94	0	0	0	1	0	0	0	1	0	99	Alveolata(100)	Dinophyta(100)	Dinophyceae(100)	Dinophyceae_X(100)	Suessiales(100)	unclassified	unclassified
OTU144	71	4	3	Brackish	19	24	6	22	0	2	1	1	2	0	1	0	78	Stramenopiles(100)	Stramenopiles_X(100)	Dictyochophyceae(100)	Dictyochophyceae_X(100)	Pedinellales(100)	Pedinellales_X(100)	Pedinellales_X+sp.(100)
OTU155	47	15	11	Brackish	5	1	35	6	0	0	13	2	0	0	11	0	73	Rhizaria(100)	Cercozoa(100)	Filosa-Thecofilosea(100)	Ebriida(100)	TAGIRI1-lineage(100)	TAGIRI1-lineage_X(100)	TAGIRI1-lineage_X+sp.(100)
OTU159	68	1	1	Brackish	43	0	22	3	0	0	0	1	1	0	0	0	70	Alveolata(100)	Ciliophora(100)	Litostomatea(100)	Cyclotrichia(100)	Cyclotrichia_X(100)	Cyclotrichia_XX(100)	Cyclotrichia_XX+sp.(100)
OTU167	59	3	2	Brackish	0	0	58	1	0	0	3	0	0	0	2	0	64	Hacrobia(100)	Cryptophyta(100)	Cryptophyceae(100)	Cryptophyceae_X(100)	Goniomonadales(100)	Goniomonas(100)	Goniomonas+amphinema(100)
OTU168	61	1	1	Brackish	28	26	7	0	0	1	0	0	1	0	0	0	63	Opisthokonta(100)	Choanoflagellida(100)	Choanoflagellatea(100)	Acanthoecida(100)	Stephanoecidae_Group_D(100)	Stephanoecidae_Group_D_X(100)	Stephanoecidae_Group_D_X+sp.(100)
OTU172	48	9	3	Brackish	45	2	1	0	9	0	0	0	3	0	0	0	60	Opisthokonta(100)	Metazoa(100)	Ctenophora(100)	Ctenophora_X(100)	Ctenophora_XX(100)	unclassified(100)	unclassified(100)
OTU187	39	9	5	Brackish	0	0	39	0	0	0	9	0	0	0	5	0	53	Hacrobia(100)	Haptophyta(100)	Pavlovophyceae(100)	Pavlovales(100)	Pavlovaceae(100)	unclassified(100)	unclassified(100)
OTU197	42	2	2	Brackish	0	0	33	9	0	0	1	1	0	0	2	0	46	Archaeplastida(100)	Chlorophyta(100)	Pyramimonadales(100)	Pyramimonadales_X(100)	Pyramimonadales_XX(100)	Pyramimonadales_XXX(100)	Pyramimonadales_XXX+sp.(100)
OTU206	43	1	1	Brackish	23	9	9	2	1	0	0	0	1	0	0	0	45	Alveolata(100)	Ciliophora(100)	Spirotrichea(100)	Oligotrichia(100)	Oligotrichia_X(100)	Oligotrichia_XX(100)	Oligotrichia_XX+sp.(100)
OTU212	34	4	2	Brackish	0	34	0	0	0	1	0	3	1	0	1	0	40	Alveolata(100)	Ciliophora(100)	Litostomatea(100)	Haptoria(100)	unclassified(100)	unclassified(100)	unclassified(100)
OTU222	35	1	1	Brackish	31	0	0	4	0	0	0	1	1	0	0	0	37	Rhizaria(100)	Cercozoa(100)	Filosa-Thecofilosea(100)	Ventricleftida(100)	CCW10-lineage(100)	CCW10-lineage_X(100)	CCW10-lineage_X+sp.(100)
OTU226	31	3	1	Brackish	22	1	5	3	2	0	1	0	1	0	0	0	35	Alveolata(100)	Ciliophora(100)	Spirotrichea(100)	Choreotrichia(100)	Strobilidiidae(100)	Strobilidium(100)	Strobilidium+sp.(100)
OTU233	34	1	1	Brackish	15	2	0	17	0	0	0	1	0	1	0	0	36	Hacrobia(100)	Picobiliphyta(100)	Picobiliphyta_X(100)	Picobiliphyta_XX(100)	Picobiliphyta_XXX(100)	Picobiliphyta_XXXX(100)	Picobiliphyta_XXXX+sp.(100)
OTU236	31	3	1	Brackish	1	27	3	0	1	2	0	0	1	0	0	0	35	Alveolata(100)	Dinophyta(100)	Syndiniales(100)	Dino-Group-II(100)	unclassified(100)	unclassified(100)	unclassified(100)
OTU243	25	4	3	Brackish	14	0	2	9	1	1	0	2	3	0	0	0	32	Stramenopiles(100)	Stramenopiles_X(100)	Eustigmatophyceae(100)	Eustigmatophyceae_X(100)	Eustigmatophyceae_XX(100)	Nannochloropsis(100)	unclassified(100)
OTU250	26	2	2	Brackish	10	5	3	8	0	1	0	1	0	0	2	0	30	Opisthokonta(100)	Choanoflagellida(100)	Choanoflagellatea(100)	Choanoflagellatea_X(100)	Choanoflagellatea_X_Group_L(100)	Choanoflagellatea_X_Group_L_X(100)	Choanoflagellatea_X_Group_L_X+sp.(100)
OTU260	26	2	1	Brackish	0	0	25	1	0	0	2	0	1	0	0	0	29	Rhizaria(100)	Cercozoa(100)	Filosa-Thecofilosea(100)	Cryomonadida(100)	Protaspa-lineage(100)	Protaspa-lineage_X(99)	Protaspa-lineage_X+sp.(99)
OTU271	23	5	2	Brackish	21	0	2	0	3	0	2	0	2	0	0	0	30	Archaeplastida(100)	Chlorophyta(100)	Ulvophyceae(100)	Ulotrichales(100)	Ulotrichales_X(100)	Urospora(90)	unclassified(90)
OTU281	23	2	2	Brackish	0	0	17	6	0	0	2	0	0	0	2	0	27	Rhizaria(100)	Cercozoa(100)	unclassified(100)	unclassified(100)	unclassified(100)	unclassified(100)	unclassified(100)
OTU282	24	1	1	Brackish	9	14	1	0	0	0	1	0	0	0	1	0	26	Rhizaria(100)	Cercozoa(100)	Cercozoa_X(100)	Cercozoa_XX(100)	Cercozoa_XXX(100)	Cercozoa_XXXX(100)	Cercozoa_XXXX+sp.(100)
OTU289	23	2	1	Brackish	5	10	8	0	1	1	0	0	1	0	0	0	26	Opisthokonta(100)	Metazoa(100)	unclassified(100)	unclassified(100)	unclassified(100)	unclassified(100)	unclassified(100)
OTU293	22	1	1	Brackish	14	4	4	0	1	0	0	0	1	0	0	0	24	Archaeplastida(100)	Chlorophyta(100)	Trebouxiophyceae(100)	Trebouxiophyceae_X(100)	Trebouxiophyceae_XX(100)	Choricystis(100)	Choricystis+sp.(100)
OTU332	16	3	1	Brackish	12	1	3	0	0	0	3	0	0	0	1	0	20	Rhizaria(100)	Cercozoa(100)	Filosa-Thecofilosea(100)	Cryomonadida(100)	Cryothecomonas-lineage(100)	Cryothecomonas(80)	Cryothecomonas+sp.(80)
OTU351	13	3	2	Brackish	13	0	0	0	0	3	0	0	0	2	0	0	18	Rhizaria(100)	Cercozoa(100)	Filosa-Imbricatea(100)	Filosa-Imbricatea_X(100)	Novel-clade-2(100)	Novel-clade-2_X(100)	Novel-clade-2_X+sp.(100)
OTU406	10	4	2	Brackish	7	0	3	0	4	0	0	0	2	0	0	0	16	Archaeplastida(100)	Chlorophyta(100)	Ulvophyceae(100)	Ulvales-relatives(100)	Ulvales-relatives_X(100)	Dilabifilum(88)	unclassified(81)
OTU415	9	3	1	Brackish	0	0	0	9	0	0	3	0	0	0	1	0	13	Alveolata(100)	Dinophyta(100)	Dinophyceae(100)	Dinophyceae_X(100)	unclassified(100)	unclassified(100)	unclassified(100)
OTU258	16	5	10	Brackish/Freshwater	14	0	2	0	5	0	0	0	10	0	0	0	31	Archaeplastida(100)	Chlorophyta(100)	Ulvophyceae(100)	Ulvales-relatives(100)	Ulvales-relatives_X(100)	Ulvales-relatives_XX(100)	Ulvales-relatives_XX+sp.(100)
OTU20	288	23	319	Brackish/Freshwater	203	3	20	62	4	1	4	14	315	0	3	1	630	Stramenopiles(100)	Stramenopiles_X(100)	Chrysophyceae-Synurophyceae(100)	Chrysophyceae-Synurophyceae_X(100)	Clade-C(100)	unclassified	unclassified
OTU245	20	12	6	Brackish/Mixing Zone	11	4	2	3	0	2	0	10	2	0	1	3	38	Opisthokonta(100)	Choanoflagellida(100)	Choanoflagellatea(100)	Choanoflagellatea_X(100)	Choanoflagellatea_X_Group_L(100)	Choanoflagellatea_X_Group_L_X(100)	Choanoflagellatea_X_Group_L_X+sp.(100)
OTU10	593	468	2	Brackish/Mixing Zone	14	528	39	12	2	332	1	133	2	0	0	0	1063	Stramenopiles(100)	Stramenopiles_X(100)	Bacillariophyta(100)	Bacillariophyta_X(100)	Polar-centric-Mediophyceae(100)	Skeletonema(100)	unclassified(88)
OTU31	224	100	89	Brackish/Mixing Zone	28	29	17	150	64	15	13	8	57	23	9	0	413	Rhizaria(100)	Cercozoa(100)	Filosa-Thecofilosea(100)	Cryomonadida(100)	Protaspa-lineage(100)	Protaspa-lineage_X(100)	Protaspa-lineage_X+sp.(100)
OTU32	280	125	12	Brackish/Mixing Zone	101	1	32	146	10	0	1	114	12	0	0	0	417	Alveolata(100)	Ciliophora(100)	Spirotrichea(100)	Oligotrichia(100)	Strombidiidae(100)	Strombidiidae_X(99)	Strombidiidae_X+sp.(99)
OTU119	37	48	23	Brackish/Mixing Zone	18	10	8	1	10	17	14	7	14	4	5	0	108	Hacrobia(100)	Katablepharidophyta(100)	Katablepharidaceae(100)	Katablepharidales(100)	Katablepharidales_X(100)	unclassified(100)	unclassified(100)
OTU303	6	7	3	Brackish/Mixing Zone	0	1	3	2	0	4	3	0	0	0	3	0	16	Opisthokonta(100)	Metazoa(100)	Rotifera(100)	Rotifera_X(100)	Rotifera_XX(100)	unclassified(100)	unclassified(100)
OTU432	5	5	1	Brackish/Mixing Zone	1	1	3	0	2	2	1	0	0	0	1	0	11	Hacrobia(100)	Katablepharidophyta(100)	Katablepharidaceae(100)	Katablepharidales(100)	Katablepharidales_X(100)	unclassified(83)	unclassified(83)
OTU58	128	48	76	Present at all sites	1	0	81	46	13	1	26	8	16	1	14	45	252	Stramenopiles(100)	Stramenopiles_X(100)	Chrysophyceae-Synurophyceae(100)	Chrysophyceae-Synurophyceae_X(100)	Clade-F(100)	Clade-F_X(87)	Clade-F_X+sp.(87)
OTU54	174	66	46	Present at all sites	25	2	3	144	58	0	5	3	44	0	2	0	286	Stramenopiles(100)	Stramenopiles_X(100)	Dictyochophyceae(100)	Dictyochophyceae_X(100)	Pedinellales(100)	Pedinellales_X(100)	Pedinellales_X+sp.(100)
OTU69	24	17	19	Present at all sites	1	1	17	5	1	5	5	6	3	8	7	1	60	Hacrobia(100)	Katablepharidophyta(100)	Katablepharidaceae(100)	Katablepharidales(100)	Katablepharidales_X(100)	Katablepharidales_XX(100)	Katablepharidales_XX+sp.(100)
OTU74	*62*	*79*	*67*	*Present at all sites*	*27*	*1*	*3*	*31*	*39*	*6*	*13*	*21*	*34*	*3*	*9*	*21*	*208*	*Stramenopiles(100)*	*Stramenopiles_X(100)*	*Chrysophyceae-Synurophyceae(100)*	*Chrysophyceae-Synurophyceae_X(100)*	*Clade-H(100)*	*Clade-H_X(100)*	*Clade-H_X+sp.(100)*
OTU122	33	18	25	Present at all sites	28	1	4	0	3	1	6	8	2	1	17	5	76	Archaeplastida(100)	Chlorophyta(100)	Chlorophyceae(100)	Chlorophyceae_X(100)	Sphaeropleales(100)	unclassified	unclassified
OTU170	29	18	13	Present at all sites	21	0	4	4	7	2	6	3	12	1	0	0	60	Stramenopiles(100)	Stramenopiles_X(100)	Bacillariophyta(100)	Bacillariophyta_X(100)	Raphid-pennate(100)	Entomoneis(95)	Entomoneis+sp.(95)
OTU182	21	21	14	Present at all sites	11	8	2	0	5	7	8	1	10	0	4	0	56	Hacrobia(100)	Katablepharidophyta(100)	Katablepharidaceae(100)	Katablepharidales(100)	Katablepharidales_X(100)	unclassified(100)	unclassified(100)
OTU51	*87*	*117*	*88*	*Present at all sites*	*0*	*0*	*55*	*32*	*33*	*20*	*38*	*26*	*32*	*8*	*44*	*4*	*292*	*Stramenopiles(100)*	*Stramenopiles_X(100)*	*MAST(100)*	*unclassified(100)*	*unclassified(100)*	*unclassified(100)*	*unclassified(100)*
OTU91	*48*	*72*	*48*	*Present at all sites*	*29*	*4*	*5*	*10*	*40*	*5*	*24*	*3*	*42*	*1*	*5*	*0*	*168*	*Stramenopiles(100)*	*Stramenopiles_X(100)*	*Bolidophyceae-and-relatives(100)*	*Bolidophyceae-and-relatives_X(100)*	*Bolidophyceae-and-relatives_XX(100)*	*Bolidophyceae-and-relatives_XXX(100)*	*Bolidophyceae-and-relatives_XXX+sp.(100)*
OTU96	6	22	129	Freshwater	0	0	1	5	6	1	1	14	5	1	1	122	157	Stramenopiles(100)	Stramenopiles_X(100)	Chrysophyceae-Synurophyceae(100)	Synurales(100)	Synurales_X(100)	Synurales_XX(100)	Synurales_XX+sp.(100)
OTU38	67	97	206	Freshwater	0	5	57	5	54	1	7	35	66	2	10	128	370	Stramenopiles(100)	Stramenopiles_X(100)	Chrysophyceae-Synurophyceae(100)	Chrysophyceae-Synurophyceae_X(100)	Clade-C(100)	Clade-C_X(100)	Clade-C_X+sp.(100)
OTU62	34	61	149	Freshwater	13	5	6	10	25	1	10	25	37	4	21	87	244	Hacrobia(100)	Cryptophyta(100)	Cryptophyceae(100)	Cryptophyceae_X(100)	Basal_Cryptophyceae-1(100)	Basal_Cryptophyceae-1_X(100)	Basal_Cryptophyceae-1_X+sp.(100)
OTU66	1	76	159	Freshwater	0	1	0	0	5	42	2	27	10	140	4	5	236	Alveolata(100)	Ciliophora(100)	Spirotrichea(100)	Choreotrichia(100)	Choreotrichia_X(100)	Choreotrichia_XX(100)	Choreotrichia_XX+sp.(100)
OTU71	1	58	147	Freshwater	0	1	0	0	15	18	12	13	43	47	16	41	206	Alveolata(100)	Ciliophora(100)	Spirotrichea(100)	Choreotrichia(100)	Strobilidiidae(100)	Strobilidiidae_X(100)	Strobilidiidae_X+sp.(100)
OTU94	1	43	120	Freshwater	0	1	0	0	22	3	8	10	65	13	26	16	164	Alveolata(100)	Ciliophora(100)	Spirotrichea(100)	Choreotrichia(100)	Strobilidiidae(100)	Strobilidiidae_X(100)	Strobilidiidae_X+sp.(100)
OTU123	1	22	81	Freshwater	0	1	0	0	13	4	3	2	13	4	2	62	104	Stramenopiles(100)	Stramenopiles_X(100)	Chrysophyceae-Synurophyceae(100)	Chrysophyceae-Synurophyceae_X(100)	Clade-C(100)	Clade-C_X(99)	Clade-C_X+sp.(99)
OTU186	7	10	33	Freshwater	4	1	1	1	3	0	1	6	9	0	4	20	50	Hacrobia(100)	Cryptophyta(100)	Cryptophyceae(100)	Cryptophyceae_X(100)	Basal_Cryptophyceae-1(100)	Basal_Cryptophyceae-1_X(100)	Basal_Cryptophyceae-1_X+sp.(100)
OTU214	2	9	27	Freshwater	0	0	0	2	0	0	9	0	0	0	27	0	38	Opisthokonta(100)	Metazoa(100)	Rotifera(100)	Rotifera_X(100)	Rotifera_XX(100)	unclassified(100)	unclassified(100)
OTU17	1	260	420	Freshwater/Mixing Zone	0	1	0	0	140	8	67	45	127	64	212	17	681	Hacrobia(100)	Cryptophyta(100)	Cryptophyceae(100)	Cryptophyceae_X(100)	Cryptomonadales(100)	Cryptomonas(100)	Cryptomonas+sp.(96)
OTU67	2	83	132	Freshwater/Mixing Zone	0	1	0	1	0	11	45	27	1	7	75	49	217	Stramenopiles(100)	Stramenopiles_X(100)	Bacillariophyta(100)	Bacillariophyta_X(100)	Raphid-pennate(100)	Nitzschia(100)	Nitzschia+acicularis(100)
OTU1	***689***	***11031***	***8556***	***Freshwater/Mixing Zone***	***14***	***657***	***5***	***13***	***2859***	***4433***	***1082***	***2657***	***2339***	***4833***	***898***	***486***	***20276***	***Stramenopiles(100)***	***Stramenopiles_X(100)***	***Bacillariophyta(100)***	***Bacillariophyta_X(100)***	***Polar-centric-Mediophyceae(100)***	***Stephanodiscus(100)***	***Stephanodiscus+minutulus(93)***
OTU2	***103***	***1767***	***2345***	***Freshwater/Mixing Zone***	***5***	***6***	***10***	***82***	***21***	***30***	***839***	***877***	***21***	***27***	***1771***	***526***	***4215***	***Stramenopiles(100)***	***Stramenopiles_X(100)***	***Bacillariophyta(100)***	***Bacillariophyta_X(100)***	***Polar-centric-Mediophyceae(100)***	***Cyclotella(100)***	***unclassified***
OTU3	***139***	***448***	***975***	***Freshwater/Mixing Zone***	***2***	***3***	***1***	***133***	***1***	***2***	***221***	***224***	***1***	***3***	***310***	***661***	***1562***	***Stramenopiles(100)***	***Stramenopiles_X(100)***	***Bacillariophyta(100)***	***Bacillariophyta_X(100)***	***Polar-centric-Mediophyceae(100)***	***Skeletonema(100)***	***unclassified(100)***
OTU11	2	480	594	Freshwater/Mixing Zone	0	2	0	0	385	1	86	8	411	1	102	80	1076	Stramenopiles(100)	Stramenopiles_X(100)	Chrysophyceae-Synurophyceae(100)	Chrysophyceae-Synurophyceae_X(100)	Clade-F(90)	unclassified	unclassified
OTU15	2	247	279	Freshwater/Mixing Zone	1	1	0	0	185	7	44	11	224	9	27	19	528	Stramenopiles(100)	Stramenopiles_X(100)	Chrysophyceae-Synurophyceae(100)	Chrysophyceae-Synurophyceae_X(100)	Clade-C(100)	Clade-C_X(99)	Clade-C_X+sp.(99)
OTU21	8	337	257	Freshwater/Mixing Zone	0	8	0	0	136	91	59	51	138	79	27	13	602	Stramenopiles(100)	Stramenopiles_X(100)	Chrysophyceae-Synurophyceae(100)	Chrysophyceae-Synurophyceae_X(100)	Clade-E(100)	Chrysosaccus(88)	Chrysosaccus+sp.(88)
OTU26	*31*	*175*	*208*	*Freshwater/Mixing Zone*	*1*	*4*	*0*	*26*	*31*	*52*	*26*	*66*	*24*	*59*	*54*	*71*	*414*	*Stramenopiles(100)*	*Stramenopiles_X(100)*	*Bacillariophyta(100)*	*Bacillariophyta_X(100)*	*Polar-centric-Mediophyceae(100)*	*unclassified*	*unclassified*
OTU44	26	59	63	Freshwater/Mixing Zone	0	26	0	0	44	4	8	3	47	1	13	2	148	Stramenopiles(100)	Stramenopiles_X(100)	Chrysophyceae-Synurophyceae(100)	Chrysophyceae-Synurophyceae_X(100)	Clade-F(100)	Clade-F_X(100)	Clade-F_X+sp.(100)
OTU50	2	132	162	Freshwater/Mixing Zone	0	2	0	0	110	13	5	4	128	12	20	2	296	Rhizaria(100)	Cercozoa(100)	Filosa-Thecofilosea(100)	Cryomonadida(100)	unclassified(92)	unclassified(92)	unclassified(92)
OTU59	1	159	118	Freshwater/Mixing Zone	0	0	0	1	0	0	159	0	0	0	117	1	278	Rhizaria(100)	Cercozoa(100)	Filosa-Imbricatea(100)	Filosa-Imbricatea_X(100)	Novel-clade-2(100)	Novel-clade-2_X(100)	Novel-clade-2_X+sp.(100)
OTU77	7	101	91	Freshwater/Mixing Zone	1	6	0	0	44	29	18	10	32	53	6	0	199	Archaeplastida(100)	Chlorophyta(100)	Chlorophyceae(100)	Chlorophyceae_X(100)	CW-Chlamydomonadales(100)	Chlamydomonas(97)	unclassified(90)
OTU80	1	103	90	Freshwater/Mixing Zone	0	1	0	0	75	1	25	2	67	5	18	0	194	Alveolata(100)	Dinophyta(100)	Dinophyceae(100)	Dinophyceae_X(100)	Dinophyceae_XX(93)	Woloszynskia(90)	Woloszynskia+pascheri(90)
OTU88	1	81	90	Freshwater/Mixing Zone	0	1	0	0	51	0	17	13	53	0	15	22	172	Stramenopiles(100)	Stramenopiles_X(100)	Chrysophyceae-Synurophyceae(100)	Chrysophyceae-Synurophyceae_X(100)	unclassified	unclassified	unclassified
OTU89	1	101	57	Freshwater/Mixing Zone	0	0	0	1	7	29	3	62	4	21	4	28	159	Hacrobia(100)	Katablepharidophyta(100)	Katablepharidaceae(100)	Katablepharidales(100)	Katablepharidales_X(100)	Katablepharidales_XX(100)	Katablepharidales_XX+sp.(100)
OTU101	7	59	77	Freshwater/Mixing Zone	5	1	0	1	30	4	7	18	35	2	4	36	143	Stramenopiles(100)	Stramenopiles_X(100)	Chrysophyceae-Synurophyceae(100)	Chrysophyceae-Synurophyceae_X(100)	Clade-C(100)	unclassified(100)	unclassified(100)
OTU110	27	37	55	Freshwater/Mixing Zone	2	17	6	2	30	2	4	1	39	4	10	2	119	Stramenopiles(100)	Stramenopiles_X(100)	Chrysophyceae-Synurophyceae(100)	Chrysophyceae-Synurophyceae_X(100)	Clade-F(100)	unclassified(93)	unclassified(93)
OTU115	1	47	63	Freshwater/Mixing Zone	0	1	0	0	28	6	2	11	22	31	9	1	111	Archaeplastida(100)	Chlorophyta(100)	Trebouxiophyceae(100)	Chlorellales(100)	Chlorellales_X(100)	unclassified(98)	unclassified(98)
OTU118	3	56	51	Freshwater/Mixing Zone	0	2	0	1	13	24	8	11	11	32	5	3	110	Stramenopiles(100)	Stramenopiles_X(100)	Bacillariophyta(100)	Bacillariophyta_X(100)	Raphid-pennate(100)	unclassified	unclassified
OTU134	1	46	42	Freshwater/Mixing Zone	0	1	0	0	34	4	8	0	32	1	9	0	89	Rhizaria(100)	Cercozoa(100)	Filosa-Imbricatea(100)	Filosa-Imbricatea_X(100)	Novel-clade-2(100)	Novel-clade-2_X(100)	Novel-clade-2_X+sp.(100)
OTU141	3	30	46	Freshwater/Mixing Zone	0	0	0	3	20	5	1	4	9	34	2	1	79	Archaeplastida(100)	Chlorophyta(100)	Chlorophyceae(100)	Chlorophyceae_X(100)	CW-Chlamydomonadales(100)	Chlamydomonas(88)	unclassified(85)
OTU165	4	39	21	Freshwater/Mixing Zone	0	4	0	0	0	34	0	5	0	21	0	0	64	Opisthokonta(100)	Fungi(100)	Chytridiomycota(100)	Chytridiomycotina(100)	Chytridiomycetes(100)	Rhizophydium(100)	Rhizophydium+sp.(100)
OTU175	1	20	32	Freshwater/Mixing Zone	0	0	0	1	18	0	0	2	15	3	13	1	53	Hacrobia(100)	Cryptophyta(100)	Cryptophyceae(100)	Cryptophyceae_X(100)	Cryptomonadales(100)	Cryptomonas(100)	unclassified
OTU191	10	24	20	Freshwater/Mixing Zone	3	2	0	5	3	2	2	17	4	2	0	14	54	Opisthokonta(100)	Opisthokonta_X(100)	Fonticulea(100)	Fonticulida(100)	Fonticulidae(100)	Fonticulidae_X(100)	Fonticulidae_X+sp.(100)
OTU198	2	23	19	Freshwater/Mixing Zone	0	2	0	0	10	9	3	1	10	6	2	1	44	Stramenopiles(100)	Stramenopiles_X(100)	Bacillariophyta(100)	Bacillariophyta_X(100)	Araphid-pennate(100)	Synedra(100)	unclassified(93)
OTU213	1	21	20	Freshwater/Mixing Zone	0	1	0	0	8	6	7	0	5	6	9	0	42	Stramenopiles(100)	Stramenopiles_X(100)	Bacillariophyta(100)	Bacillariophyta_X(100)	Araphid-pennate(100)	Fragilaria(100)	Fragilaria+nanana(100)
OTU295	3	11	11	Freshwater/Mixing Zone	0	0	0	3	0	0	9	2	0	0	4	7	25	Stramenopiles(100)	Stramenopiles_X(100)	Labyrinthulea(100)	Thraustochytriales(100)	Thraustochytriaceae(100)	unclassified	unclassified
OTU313	1	4	5	Freshwater/Mixing Zone	0	0	1	0	0	1	0	3	0	5	0	0	10	Hacrobia(100)	Katablepharidophyta(100)	Katablepharidaceae(100)	Katablepharidales(100)	Katablepharidales_X(100)	Katablepharidales_XX(100)	Katablepharidales_XX+sp.(100)
OTU318	5	9	7	Freshwater/Mixing Zone	0	1	3	1	5	4	0	0	3	2	2	0	21	Rhizaria(100)	Cercozoa(100)	Filosa-Thecofilosea(100)	Ebriida(100)	Botuliformidae(100)	Botuliformidae_X(100)	Botuliformidae_X+sp.(100)
OTU329	2	7	9	Freshwater/Mixing Zone	0	0	0	2	0	0	6	1	0	0	8	1	18	Stramenopiles(100)	Stramenopiles_X(100)	Oomyceta(100)	Oomyceta_X(100)	Oomyceta_XX(100)	unclassified(87)	unclassified(87)

OTU420	1	6	6	Freshwater/Mixing Zone	0	1	0	0	5	1	0	0	2	2	2	0	13	Hacrobia(100)	Haptophyta(100)	Prymnesiophyceae(100)	unclassified(100)	unclassified(100)	unclassified(100)	unclassified(100)
OTU443	1	6	6	Freshwater/Mixing Zone	1	0	0	0	3	3	0	0	0	6	0	0	13	Archaeplastida(100)	Chlorophyta(100)	Chlorophyceae(100)	Chlorophyceae_X(100)	CW-Chlamydomonadales(100)	Pascherina(100)	Pascherina+tetras(100)
OTU456	1	4	6	Freshwater/Mixing Zone	0	1	0	0	3	0	1	0	5	0	1	0	11	Rhizaria(100)	Cercozoa(100)	unclassified(100)	unclassified(100)	unclassified(100)	unclassified(100)	unclassified(100)
OTU82	1	128	24	Mixing Zone	0	1	0	0	1	9	115	3	1	4	18	1	153	Stramenopiles(100)	Stramenopiles_X(100)	Chrysophyceae-Synurophyceae(100)	Chrysophyceae-Synurophyceae_X(100)	unclassified(100)	unclassified(100)	unclassified(100)
OTU83	1	104	63	Mixing Zone	1	0	0	0	15	27	22	40	14	21	6	22	168	Alveolata(100)	Ciliophora(100)	Prostomatea(100)	Prostomatea-1(100)	Prostomatea-1_X(100)	Prostomatea-1_XX(100)	Prostomatea-1_XX+sp.(100)
OTU154	2	46	25	Mixing Zone	0	2	0	0	15	22	4	5	12	12	1	0	73	Rhizaria(100)	Cercozoa(100)	Filosa-Thecofilosea(100)	Cryomonadida(100)	Protaspa-lineage(100)	Protaspa-lineage_X(100)	Protaspa-lineage_X+sp.(100)
OTU156	5	42	18	Mixing Zone	0	5	0	0	0	28	5	9	0	17	0	1	65	Opisthokonta(100)	Fungi(100)	Chytridiomycota(100)	Chytridiomycotina(100)	Chytridiomycetes(100)	Rhizophydium(98)	Rhizophydium+sp.(98)
OTU192	5	36	8	Mixing Zone	3	2	0	0	0	22	0	14	1	7	0	0	49	Stramenopiles(100)	Stramenopiles_X(100)	Chrysophyceae-Synurophyceae(100)	Chrysophyceae-Synurophyceae_X(100)	unclassified(100)	unclassified(100)	unclassified(100)
OTU199	2	39	3	Mixing Zone	0	0	0	2	2	0	0	37	1	0	0	2	44	Hacrobia(100)	Cryptophyta(100)	Cryptophyceae(100)	Cryptophyceae_X(100)	Cryptomonadales(100)	unclassified(91)	unclassified(91)
OTU305	7	13	4	Mixing Zone	0	1	6	0	0	0	13	0	0	0	4	0	24	Archaeplastida(100)	Rhodophyta(100)	Rhodellophyceae(100)	Rhodellales(100)	Rhodellales_X(100)	Rhodella(100)	unclassified(100)
OTU310	3	12	6	Mixing Zone	2	1	0	0	5	3	2	2	5	0	1	0	21	Stramenopiles(100)	Stramenopiles_X(100)	Chrysophyceae-Synurophyceae(100)	Chrysophyceae-Synurophyceae_X(100)	unclassified	unclassified	unclassified
OTU374	4	9	3	Mixing Zone	0	0	1	3	0	0	8	1	0	0	3	0	16	Stramenopiles(100)	Stramenopiles_X(100)	Labyrinthulea(100)	Thraustochytriales(100)	Thraustochytriaceae(100)	unclassified	unclassified
OTU368	6	7	3	Rare	5	0	1	0	2	4	0	1	2	1	0	0	16	Stramenopiles(100)	Stramenopiles_X(100)	Bacillariophyta(100)	Bacillariophyta_X(100)	Araphid-pennate(100)	Diatoma(100)	Diatoma+tenue(99)
OTU413	1	7	4	Rare	0	1	0	0	0	2	2	3	0	3	1	0	12	unclassified(100)	unclassified(100)	unclassified(100)	unclassified(100)	unclassified(100)	unclassified(100)	unclassified(100)
OTU475	1	6	3	Rare	0	0	0	1	0	0	6	0	0	0	3	0	10	Stramenopiles(100)	Stramenopiles_X(100)	Pirsonia(100)	Pirsonia_X(100)	Pirsonia_XX(100)	Pirsonia_XXX(100)	Pirsonia_XXX+sp.(100)
OTU201	1	2	5	Rare	0	1	0	0	0	1	0	1	0	4	1	0	8	Archaeplastida(100)	Chlorophyta(100)	Chlorophyceae(100)	Chlorophyceae_X(100)	CW-Chlamydomonadales(100)	Chlamydomonas(81)	unclassified
OTU238	1	2	4	Rare	0	0	0	1	2	0	0	0	3	0	0	1	7	Alveolata(100)	Ciliophora(100)	Spirotrichea(100)	Hypotrichia(100)	unclassified	unclassified	unclassified
OTU244	1	5	8	Rare	0	1	0	0	0	3	2	0	0	0	4	4	14	Opisthokonta(100)	Metazoa(100)	Rotifera(100)	Rotifera_X(100)	Rotifera_XX(100)	unclassified	unclassified
OTU486	3	3	2	Rare	1	0	2	0	3	0	0	0	1	0	0	1	8	Stramenopiles(100)	Stramenopiles_X(100)	Bacillariophyta(100)	Bacillariophyta_X(100)	Raphid-pennate(100)	unclassified(100)	unclassified(100)
OTU499	1	5	2	Rare	0	0	0	1	0	0	5	0	0	0	2	0	8	unclassified(100)	unclassified(100)	unclassified(100)	unclassified(100)	unclassified(100)	unclassified(100)	unclassified(100)
OTU539	3	2	2	Rare	2	0	0	1	0	0	2	0	0	0	1	1	7	Stramenopiles(100)	Stramenopiles_X(100)	Eustigmatophyceae(100)	Eustigmatophyceae_X(100)	Eustigmatophyceae_XX(100)	unclassified(100)	unclassified(100)
OTU549	1	2	5	Rare	0	0	1	0	0	0	0	2	2	1	1	1	8	Stramenopiles(100)	Stramenopiles_X(100)	Bacillariophyta(100)	Bacillariophyta_X(100)	Raphid-pennate(100)	Navicula(100)	unclassified(82)
OTU551	1	4	3	Rare	0	1	0	0	2	0	2	0	3	0	0	0	8	Alveolata(100)	Ciliophora(100)	Oligohymenophorea(100)	Sessilida(100)	Vorticellidae(100)	unclassified(95)	unclassified(95)
OTU566	4	2	1	Rare	0	0	4	0	0	0	2	0	0	0	1	0	7	Alveolata(100)	Dinophyta(100)	Dinophyceae(100)	Dinophyceae_X(100)	Dinophyceae_XX(100)	Gyrodinium(100)	Gyrodinium+dorsum(100)
OTU599	2	1	2	Rare	2	0	0	0	1	0	0	0	2	0	0	0	5	Alveolata(100)	Ciliophora(100)	Spirotrichea(100)	Choreotrichia(100)	Choreotrichia-1(100)	Choreotrichia-1_X(100)	Choreotrichia-1_X+sp.(100)
OTU642	2	2	1	Rare	1	1	0	0	0	0	2	0	0	0	1	0	5	Alveolata(100)	Ciliophora(100)	Litostomatea(100)	Cyclotrichia(100)	unclassified	unclassified	unclassified
OTU654	2	1	1	Rare	1	1	0	0	0	0	1	0	0	0	1	0	4	Stramenopiles(100)	Stramenopiles_X(100)	Pirsonia(100)	Pirsonia_X(100)	Pirsonia_XX(100)	Pirsonia_XXX(96)	Pirsonia_XXX+sp.(96)
OTU683	2	1	1	Rare	0	0	2	0	0	0	0	1	0	0	1	0	4	Rhizaria(100)	Cercozoa(100)	Filosa-Thecofilosea(100)	Filosa-Thecofilosea_X(91)	Filosa-Thecofilosea_XX(91)	Filosa-Thecofilosea_XXX(91)	Filosa-Thecofilosea_XXX+sp.(91)
OTU857	1	1	1	Rare	1	0	0	0	0	0	0	1	1	0	0	0	3	Hacrobia(100)	Cryptophyta(100)	Cryptophyceae(100)	Cryptophyceae_X(100)	Basal_Cryptophyceae-1(100)	Basal_Cryptophyceae-1_X(100)	Basal_Cryptophyceae-1_X+sp.(100)
sum	**18531**	**19503**	**17795**		**5148**	**4895**	**4719**	**3769**	**5128**	**5726**	**3396**	**5253**	**5123**	**5733**	**4276**	**2663**	**55829**							



Supplementary Figure S1. Venn diagrams showing number of unique and shared OTUs between the different basins of the Baltic Sea, based on the data from Hu et al (2016). BB – Gulf of Bothnia (salinity 2.2-5.4), BP – Baltic Proper (salinity 5.6-7.2), A – Arkona Basin (salinity 7.2-9.7), K – Kattegat (salinity 19.8-24.2).