ECC 2014
5th Early Career Scientists Conference for Marine and Climate Research 2014

ALFRED-WEGENER-INSTITUT
HELMHOLTZ-ZENTRUM FÜR POLAR- UND MEERESFORSCHUNG

K. Hardge, S. Neuhaus, S. Frickenhaus, E.S. Kilias, C. Wolf and K. Metfies

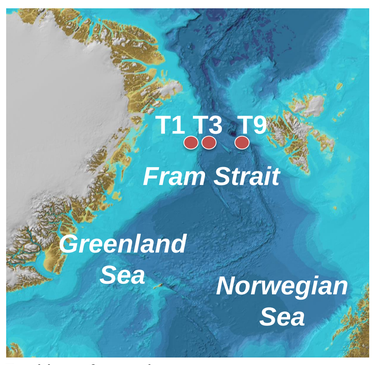
Assessment of eukaryotic communities in environmental samples:

A workflow comparison for next-generation sequencing data

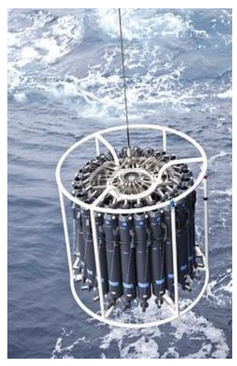
Protists are the base of food web and important primary producers in aquatic systems, such as the Arctic Ocean[1].
The composition of protist communities helps us to understand function and stability of aquatic ecosystems.
For analyzing the protist diversity, next-generation sequencing (e.g. 454 pyrosequencing) has replaced conventional methods (e.g. light microscopy). So far, there is no consensus about how to process the huge amount of sequencing data.

In order to estimate the protist composition in environmental samples as precisely as possible, this study:
(i) compares different sequence processing workflows and
(ii) combines conventional microscopy and next-generation sequencing.

Sample collection and preparation

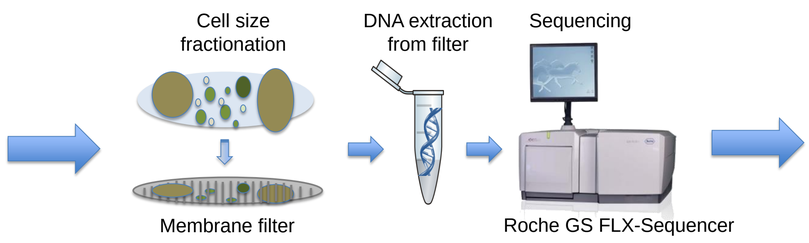


Position of samples.
Base of map: http://www.ngdc.noaa.gov/mgg/bathymetry/arctic/

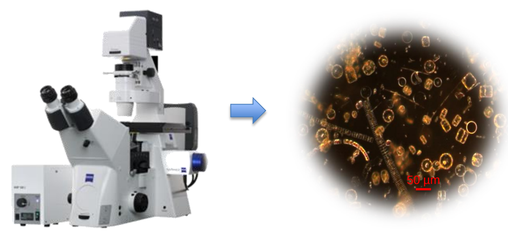


CTD-Rosette

Water samples (T1, T3, T9) were collected during *RV Polarstern* cruise to the Fram Strait with a CTD-Rosette (conductivity, temperature and depth) from the respective chlorophyll maximum layer depth (15 – 35 m) in July 2010.



DNA of protists was extracted from fractionated water samples and sequenced with 454 pyrosequencing.



A part of the clearly recognizable protist community (i.e. diatoms, belong to kingdom Stramenopila) was identified and counted using an inverted **light microscope (LM)**.

Sequence processing

The compared workflows were created with open-source software Qiime[2] **(Q)**, Mothur[3] **(M)** and PhyloAssigner[4] **(P)** by using default parameters

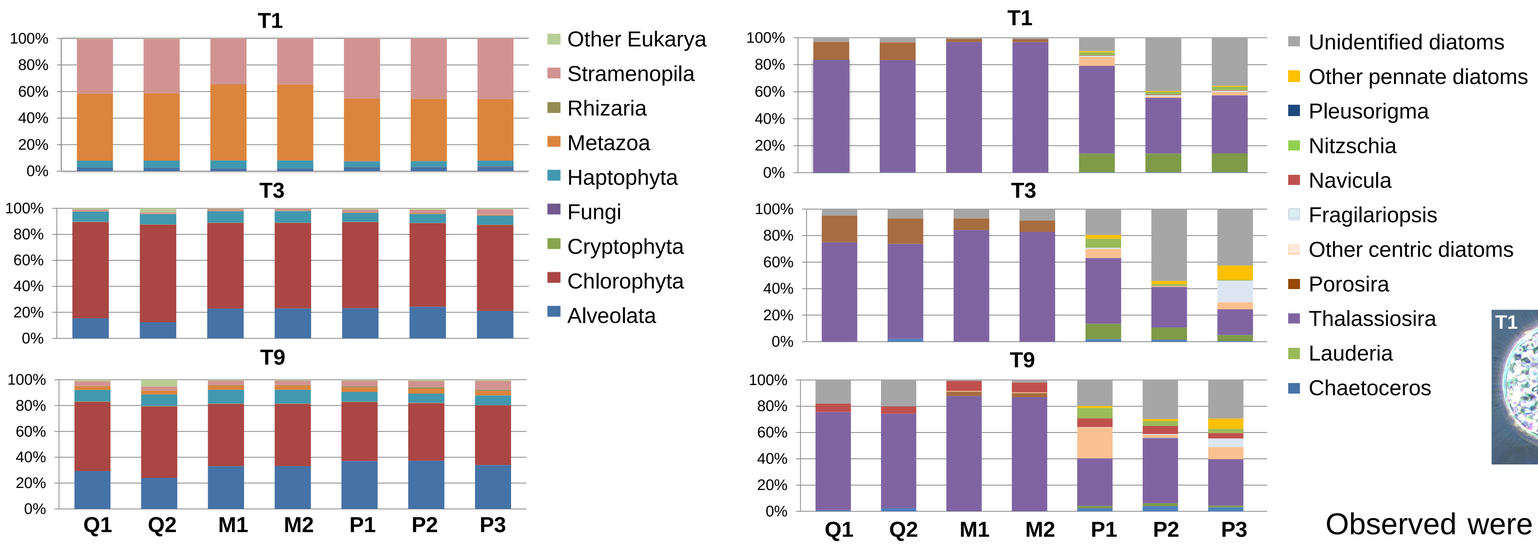
- **Quality-check:** Deletion of ambiguous bases, hybrid sequences and repeats of single bases, sequences were trimmed
- **Denoising**: Sequencing-error correction
- **Clustering:** Similar sequences are combined into operational taxonomic units (similarity threshold of 97%)

Work-flow	Quality-check	De-noising	Cluster-ing	Sequence assignment
Q1	x	-	x	Similarity based with complete reference database*
Q2	x	x	x	
M1	x	-	x	
M2	x	x	x	
P1	x	-	x	Tree based with subset of reference database*
P2	x	-	-	
P3	-	-	-	

*Silva SSU Ref NR 111

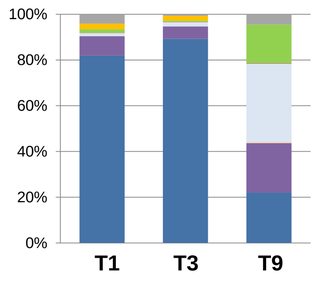
How does sequence processing effect the composition of protists?
Are the results of microscopy and sequencing comparable?

Next-generation sequencing of eukaryotic kingdoms and diatom genera



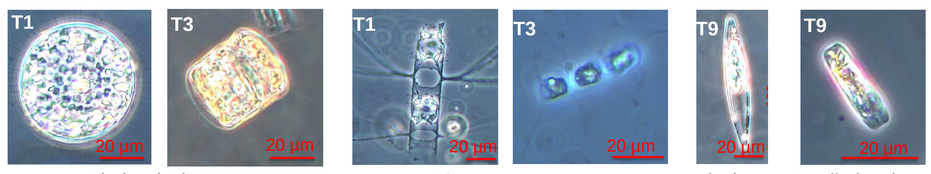
Relative sequence abundance of eukaryotic kingdoms (left) and diatom genera (right). Workflows created with PhyloAssigner (P) resulted in a higher diatom diversity. Total number of raw protist sequences (P3): T1: 41750 seqs., T3: 25407 seqs., T9: 34466 seqs.

Light microscopy of diatom genera



Relative abundance of diatom genera (same color code as used for sequencing of diatoms) and diatom cells per litre observed with microscope.

T1: 184080 Ind/L
T3: 110380 Ind/L
T9: 17040 Ind/L



Thalassiosira spp. *Chaetoceros* spp. *Navicula* sp. *Fragilariopsis* sp.

Observed were single large and healthy cells (e.g. Thalassiosira) but also chains of small, less healthy and broken cells (e.g. Chaetoceros). These cell conditions give information about succession of diatom bloom (i.e. Chaetoceros bloom prior Thalassiosira).

The effect of sequence processing

- No strong effect on kingdoms but on closer related organisms (genera).
- Default workflows of Qiime and Mothur reduced the diatom diversity (may be not appropriate for eukaryotic sequences).
- A phylogenetic placement of sequences is more reliable than a similarity based assignment (esp. for unknown species as found in the Arctic Ocean).
- Genetic similarity of > 97% is too low for determining real diatom diversity.

✓ Sequencing allowed a reproducible overview of protist kingdoms.

A comparison of conventional and molecular methods

- Results of PhyloAssigner were comparable with microscopic observations.
- Some counted diatom genera were not detected via sequencing due to degraded cell content (e.g. Chaetoceros).
- Possible misidentification occurred due to similar morphological features.
- Rare species could not be detected with microscopy (only 50 ml analyzed).

✓ Microscopy gave useful information about the diversity and ecology of dominant diatoms in the water samples.

Sequence processing parameters have to be chosen individually according to the scope of project and taxonomic level.
A combination of molecular and conventional methods provides valuable insights into the real conditions in the field and allows a better comparability between diversity studies.

References :
1 Gosselin, M. *et al.* New measurements of phytoplankton and ice algal production in the Arctic Ocean. *Deep Sea Res. Part II* **44,** 1623–1644 (1997).
2 Caporaso, G.J. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335–336 (2010).
3 Schloss, P.D. *et al.* Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **75,** 7537–7541 (2009).
4 Vergin, K.L. *et. al.* High-resolution SAR11 ecotype dynamics at the Bermuda Atlantic Time-series Study site by phylogenetic placement of pyrosequences. *ISME J.* **7**, 1322–1332 (2013).

Acknowledgements:
This study was accomplished within the Young Investigator Group PLANKTOSENS, funded by the Initiative and Networking Fund of the Helmholtz Association. We thank the captain and crew of the RV *Polarstern* for their support during the cruise. We are grateful to A. Nicolaus and K. Oetjen for technical support in the laboratory.

ALFRED-WEGENER-INSTITUT
HELMHOLTZ-ZENTRUM FÜR POLAR- UND MEERESFORSCHUNG
Kristin Hardge
Am Handelshafen 12
27570 Bremerhaven
Telefon 0471 4831-2261
www.awi.de
kristin.hardge@awi.de