Project:	BLUEPRINT
Deliverable:	5.2
Title:	Matching the BALTSEM structure to three classes of metagenomic data
Work package:	5
Deadline:	Month 30
Nature:	Model
Dissemination level:	Public
Responsible contact person:	

Åke Hagström, Linneaus University, Sweden, ake.hagstrom@lnu.se

Biogeochemical modeling is an essential tool, to be used in the design of management actions, required to reach environmental targets set e.g. by HELCOM's Baltic Sea Action Plan (BSAP) and the EU Marine Strategy Framework Directive (MSFD). It is therefore an important task to generate additional scientific support for the formulation of Baltic biogeochemical models. The goals of WP5 are to use microbial blueprints to validate and improve parameterization of basin-scale biogeochemical models of the Baltic Sea. In the Blueprint project we have chosen to refer to the BALTSEM model as a biogeochemical model prototype for the Baltic, since it is validated (Savchuk Oleg P. , 2012) and has an extensive library of model runs (Meier et al., 2012, Gustafsson, 2012).

Like other Baltic models (Eilola et al., 2011), it describes nutrient turnover in the euphotic zone by first-order mineralization of detritus and nutrient excretion, and can be illustrated in a simple graph (Fig. 1, left).

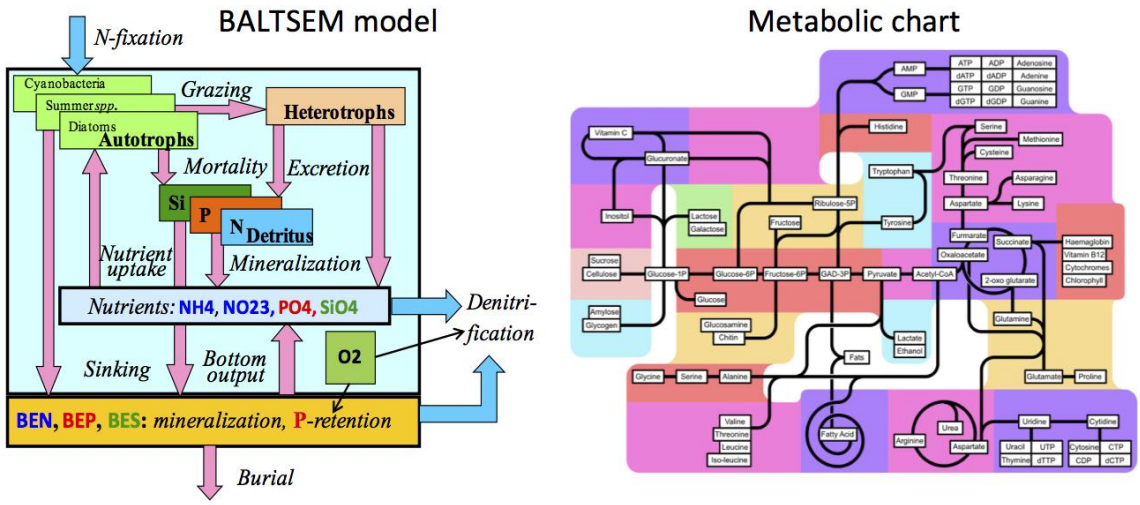

Figure 1: Comparison of overall structure between biogeochemical model formulation and outline of a metabolic chart.

BALTSEM builds on the assumption that thousands of microbial species optimize their growth rates within the constraints of the nutrient and light regime and according to their individual potential. The biogeochemical model as illustrated in Fig. 1 left, mimics the flow of matter in accordance with a conceptual ecosystem structure that links a few key compartments such as “algae” and “heterotrophs”. The complexity of linking a few components, however, quickly becomes large enough that a mathematical model is a useful tool, when analyzing different nutrient load scenarios. Due to the multitude of connections and biological core elements the outline of the model thus shows a superficial resemblance to a generic metabolic chart (Fig. 1, right).

This was the initial observation that lay behind the original description of this second deliverable of work package 5. In this deliverable we had assumed that the BALTSEM structure could be matched to three classes of metagenomic data, where the abundance of functional genes could be aligned with the flow of matter generated by the model. However, the flow of matter in the model focuses on inorganic mineral salts whereas in the biochemical pathway the flow is mainly organic matter (carbon). While these entities in an overall scale can be converted to each other by applying an approximate C:N:P ratio between the elements, the comparison between the model and the biochemistry pathway map is not as straightforward as first indicated. The model depicts transfer of matter between organisms/groups while metabolic maps show transformation of matter within the cell, through biochemical processes.

Today the entire biological system responsible for running the biogeochemistry of the sea can be observed through genetic information. This information forms the metagenome containing a number of components that is many orders of magnitude larger than those that make up the biogeochemical flows in the model. Thus, a direct comparison trying to find correlation over time between the outcome of the model and the metagenome turns out to be difficult. We therefore aim to find connecting points where the flow of matter in quantitative terms can be extracted from model runs and correlated to the abundance of molecular elements. As a result, the following modified strategies are now being investigated.

Strategy 1. Including energetics and nutrient requirements of bacterial growth

Heterotrophic microorganisms derive the energy and carbon needed for building up their biomass from a large number of organic substrates. They can cover the nitrogen and phosphorus requirement for growth either with nitrogen and phosphorus contained in the organic substrate. Most bacteria can also take up dissolved inorganic nitrogen (NH_4 and NO_3) and phosphorus (PO_4).

We have used the thermodynamic electron equivalents model, TEEM (McCarty, 2007) to derive stoichiometric relationships for microorganism growth. TEEM constructs an energy budget for growth based on the Gibbs free energies of electron donor, acceptor and biomass buildup reactions. Microbial growth stoichiometry is thus regarded as a redox reaction where the energy generated by electrons channeled into substrate catabolism matches the energy required for biomass buildup.

As an example, TEEM predicts the following stoichiometry for bacterial growth on amino acids (glycine), with oxygen as electron acceptor:

$$C(NH_3)_{0.5}OH + 0.39\ O_2 + 0.64\ H_2O + 0.009\ HPO_4^{2-} \rightarrow 0.36\ C(NH_3)_{0.2}(H_3PO4)_{0.024} + 0.43\ NH_4^+ + 0.64\ HCO_3^- + 0.19\ H^+$$

Growth stoichiometries vary with substrate energy content, its degree of carbon oxidation, and with substrate nitrogen and phosphorus content (Fig.). If bacteria grow on carbohydrates alone, mineral nitrogen and phosphorus are taken up to cover N and P requirements for biomass buildup. Peptides and amino acids contain enough nitrogen to support growth and the nitrogen surplus is excreted. For DOM composed like phytoplankton, zooplankton, heterotrophic nanoflagellates or bacterial biomass, NH_4 is always excreted whereas PO_4 is taken up for phosphorus-poor DOM (phytoplankton and zooplankton elemental composition) and released only when bacteria utilize phosphorus-rich DOM (microzooplankton and bacteria elemental ratios).

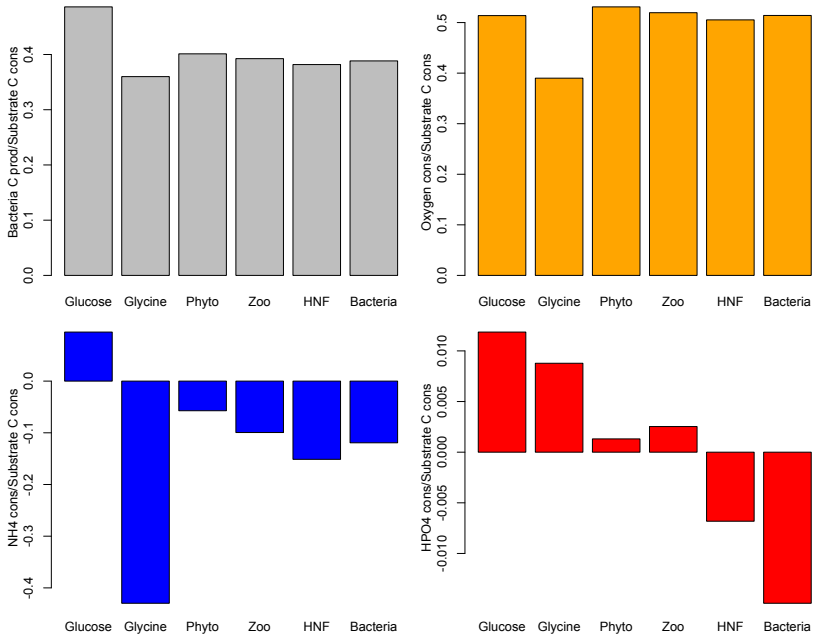

Fig. 2: Stoichiometries for bacterial growth on DOM with different composition and oxygen as electron acceptor predicted by the TEEM model. Bacterial carbon production (top left), oxygen consumption (top right), NH_4 production or consumption (bottom left) and PO_4 production or consumption (bottom right) per substrate carbon consumed.

To mimic bacteria growth on a mix of substrates, bacteria were allowed to take up organic substrates, NH_4, NO_3, and PO_4 in varying proportions so that their instantaneous growth rate is maximized. This implies that bacteria will not always utilize the most energy rich substrate; if energy poor, but nitrogen or phosphorus rich substrates are available they are taken up to supply the nitrogen and phosphorus requirements for growth.

TEEM derived stoichiometries together with the growth-rate optimization strategy were implemented into the BALTSEM model to simulate bacterial growth. To link bacterial production to the ambient concentrations of substrates and mineral nutrients, we have constrained bacterial production by a temperature dependent maximum rate and the uptake of substrates and mineral nutrients by Michaelis-Menten type kinetics. We added heterotroph nanoflagellates and ciliates to the model as grazers to simulate the fate of bacterial biomass. For the oxic part of the water column, the entire mineralization of DOM was linked to bacteria growth, replacing the first-order mineralization rates used previously.

Preliminary results showed, that bacteria growth on bulk DOM lead to high phosphate uptake and unrealistically high phosphorus limitation, strong competition with phytoplankton for phosphate and depressed phytoplankton growth. Consequently, marine DOM was split into a carbohydrate pool produced mainly by algal exudation, and a nitrogen and phosphorus rich “protein” pool. Phosphorus

taken up with “protein” DOM supplemented bacteria growth on carbohydrate DOM; consequently phosphorus uptake rates and competition with phytoplankton declined (Fig.). Still, phosphate was never excreted and phosphorus mineralization was entirely provided by micro- and mesozooplankton grazers. Both marine DOM types were preferred over terrestrial DOM, which was assumed to have a lower energy content. Seasonal dynamics of bacterial growth were simulated correctly, but slightly underestimated.

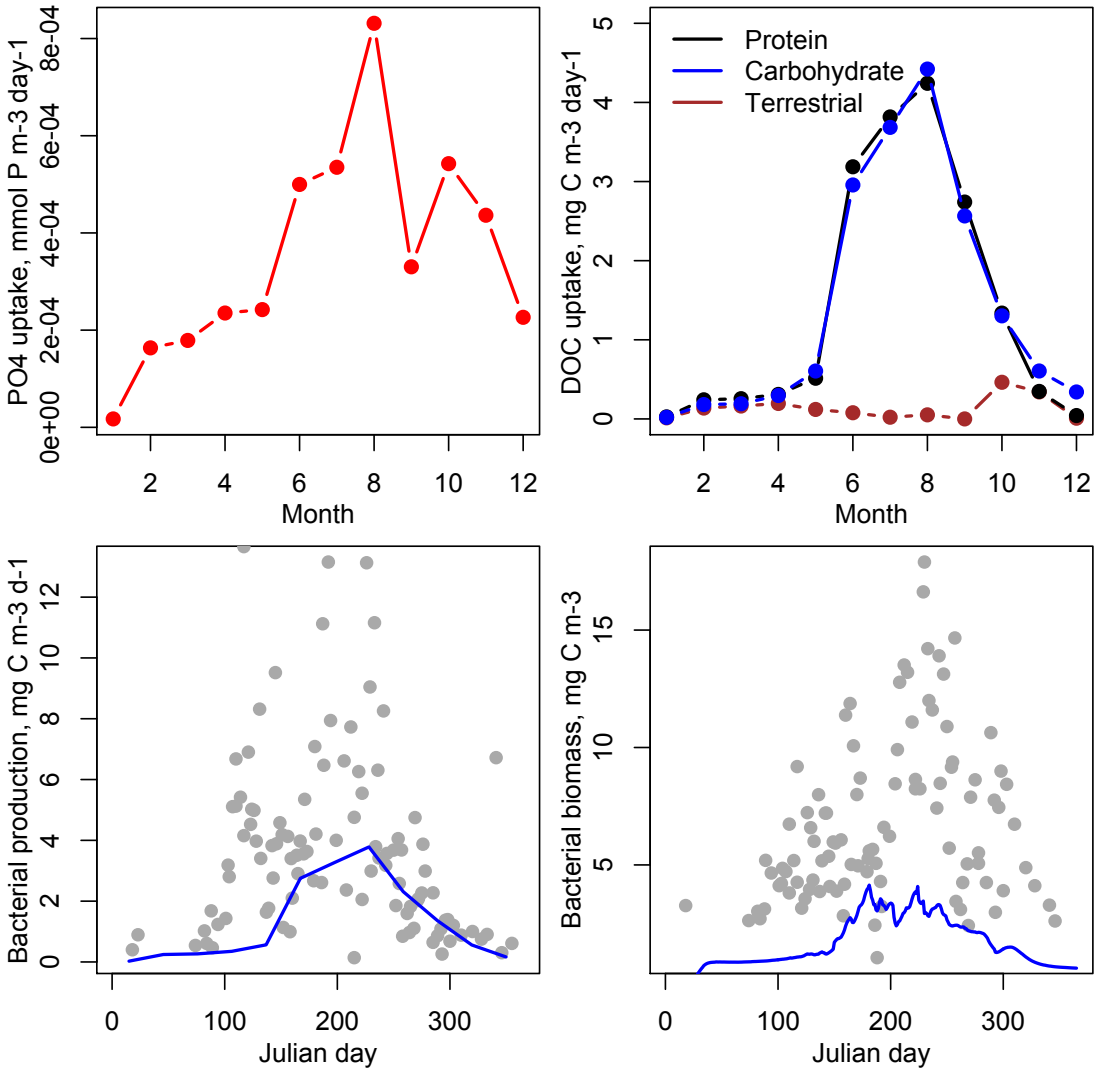

Fig. 3: Simulated bacteria growth in the Baltic Proper surface layer. Top row shows simulated phosphate (top left) and DOM uptake (top right); bottom row presents simulated bacterial production (bottom left) and bacterial biomass (bottom right) compared to observations at the Kalmar LMO station (line: simulated, dots: observed)

Strategy 2. Including transporter affinity and bacterial growth dynamics

The metagenome also shows the abundance of proteins such as transporters that guide the transfer of matter in and out of organisms. In microbial membranes these proteins may constitute about 75% by weight and often make up the majority of expressed proteins in natural communities. From the LMO 2012 time-series we have demonstrated a strong temporal signal in the distribution of different types of

transporters (John Sundh, pers. comm.). We correlated transporter abundances in both the metagenomic (DNA-level) and metatranscriptomic (mRNA-level) data to environmental parameters. Due to the seasonality at the LMO site some of these environmental parameters are strongly co-correlated (e.g. temperature/DOC and nitrate/phosphate) and thus the specific links between environment and transporter usage are difficult to discern. However, because these co-correlations are stable over several years we could identify transporters that were more abundant in either spring (being positively correlated with nitrate/phosphate) or summer (being positively correlated with temperature/DOC) and used these general patterns to describe substrate usage in the microbial community at the LMO site. Our results show that 32 out of 133 total transporters were strongly correlated with at least one of the environmental variables (absolute spearman rank correlation > 0.5) in the metagenome (Fig. 4). Of these, the ATP-driven nitrate transporter was more prevalent in spring samples while ammonium permeases and ATP-driven amino-acid transporters were more abundant in summer. In addition, the ATP-driven phosphate transporter (pst) was more abundant in summer samples (negatively correlated with phosphate levels). Also, we found ATP-driven transporters for simple sugars such as maltose and other monosaccharides to be more prevalent in summer, together with ATP-driven polysaccharide transporters.

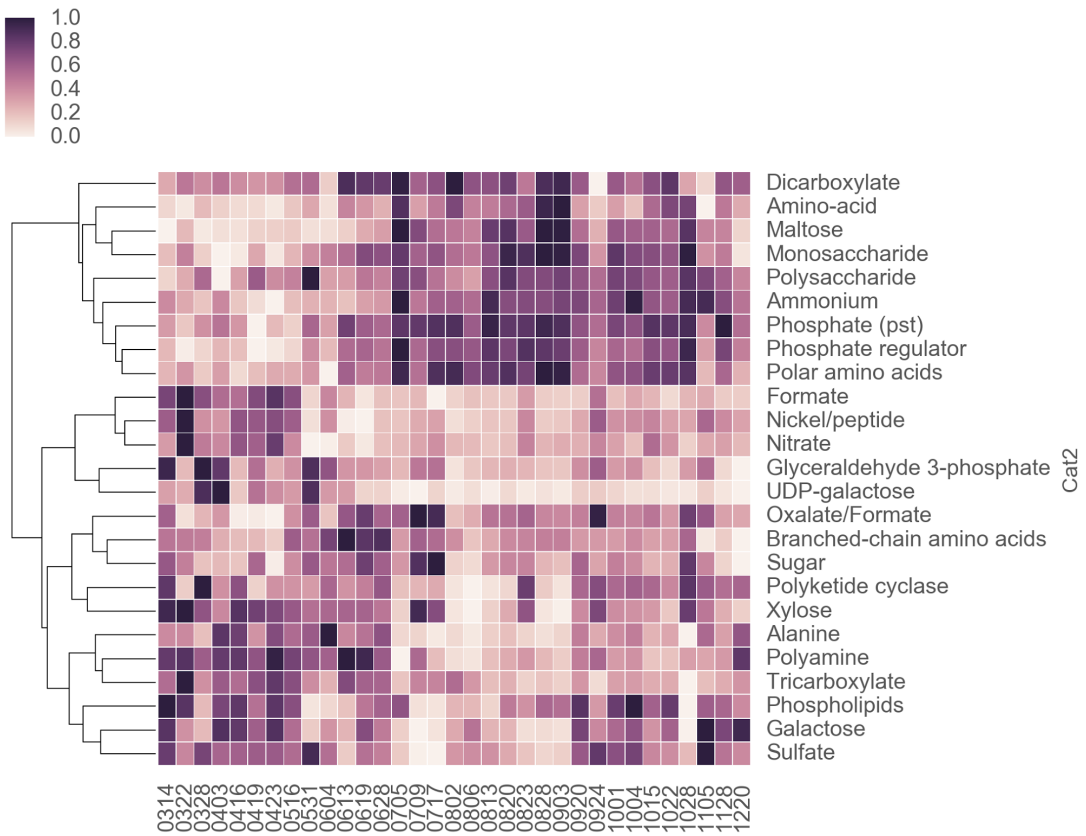

Fig. 4. Heatmap showing relative abundance of transporter substrate categories for a subset of 32 transporters strongly correlated with at least one environmental variable. Each column represents one metagenomic sample taken in 2012. Substrate categories (rows) are clustered by seasonal pattern using complete linkage clustering.

In contrast, TonB-dependent transporters known to be involved in polysaccharide transport were more abundant in spring. Further, we observed a shift in carboxylate transporters where tricarboxylate and dicarboxylate transporters of the ATP-independent TRAP-type were more abundant in spring and summer, respectively. These general patterns were also evident in the metatranscriptomic data. The

significance of these observations is that transporters are indicators of the uptake, or even export, of small molecules across cellular membranes. Since uptake and release of nutrients is an important description of the flow of matter in the BALTSEM model, the abundance of transporter genes and proteins may be an efficient way to connect metagenomic variations with model dynamics. In order to explore this argument we are currently developing a module to demonstrate behavior of different bacterial populations having different transporter affinity (see below). In the present setup, BALTSEM correctly indicates phosphate limitation for bacterial growth, which is also obvious from the abundance of phosphate ABC-type phosphate transporters in summer. DOM resolution in the model is still too coarse to represent the shift between polysaccharide and simple sugar uptake indicated by the transporter data and to correctly simulate the seasonal dynamics of nitrogen uptake.

In an attempt to illustrate the effect of transporter affinity and bacteriophage interaction a bacterial population dynamics module was constructed. The connection between transporters and bacteriophages is that transporter molecules often are attachment sites for phages. The module includes the possibility to modify a generic bacterial "species" expressing different affinity of a number of transporters related to the dominant classes of macromolecules polysaccharide carbohydrates, proteins, nucleic acids and phospholipids.

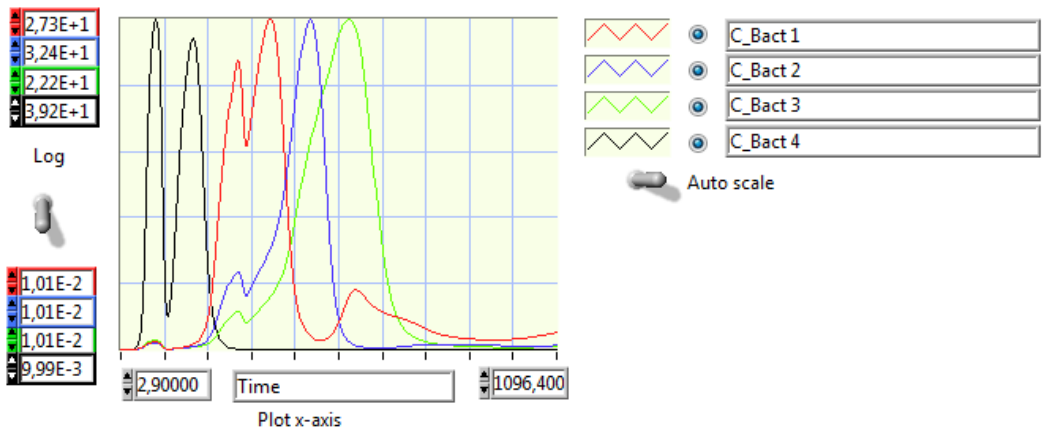

Fig. 5: Module run mimicking a pulse of organic matter triggering a succession of different dominant bacterial "species".

The population dynamics module is bottom up controlled by the supply of organic matter and top down controlled through bacteriophages, directed towards specific bacteria. Using the "killing the winner" concept (Thingstad, 2000) the success of bacteria expressing the optimal transporter configuration was counteracted by predation. The module was formulated based on two population dynamic models published in earlier work (Blackburn et al., 1996, Middelboe et al., 2001). In Fig. 5 a module run is shown where an input of primary production, mimicking the pulse of organic matter at the end of the spring bloom, triggers the succession of different dominant bacterial "species". Since the module is disconnected from further organic input the output tapers off as the organic matter is respired. The categories that we have so far are for protein, amino acids, phospholipids, carbohydrates and we assume that the bacteria are all equally keen on inorganic nitrogen and phosphorous.

The module will not be developed further until we can identify temporal trends in the metagenome data that could be used to set up a mini bacterial community with different transporter profiles. In the complete BALTSEM model the population dynamic microorganism module could then be assumed to create a transporter "blueprint" (showing a virtual temporal transporter abundance and dynamics pattern, generated by the biogeochemical model, that perhaps can be matched to real world metagenome data).

The coming year 2016 WP5 will be occupied in a continued iterative discussion on how to merge metagenome data and mathematical modeling (primarily involving WP:s 3, 4 and 5). Compared to the situation in 2015 we now have two solid strategies to investigate, which translates into a reasonably good chance of success. These strategies of how to match metagenome trends and the biogeochemical model output are also elements of the final deliverable of WP5 D5.3 aiming towards improved biogeochemical model formulations. Also through the generation of extensive datasets in the efforts to find metagenome/model correlations the development of the Metabolic Pathway Indicator (MPI) has been facilitated and has been initiated as a common task across the different work packages.

References:

BLACKBURN, N., ZWEIFEL, U. L. & HAGSTRÖM, Å. 1996. Cycling of marine dissolved organic matter. II. A model analysis. *Aquatic Microbial Ecology,* 11**,** 79-90.

EILOLA, K., GUSTAFSSON, B. G., KUZNETSOV, I., MEIER, H. E. M., NEUMANN, T. & SAVCHUK, O. P. 2011. Evaluation of biogeochemical cycles in an ensemble of three state-of-the-art numerical models of the Baltic Sea. *Journal of Marine Systems,* 88**,** 267-284.

GUSTAFSSON, B. G., SCHENK, F., BLENCKNER, T., EILOLA, K., MEIER, H.E.M., MÜLLER-KARULIS, B., NEUMANN, T., RUOHO-AIROLA, T., SAVCHUK, O.P. AND ZORITA, E. 2012. Reconstructing the development of Baltic Sea eutrophication 1850–2006. . *Ambio,* 41**,** 534-548.

JOHN SUNDH, C. B., ANDERS ANDERSSON, BÄRBEL MÜLLER-KARULIS, CHRISTOFFER KARLSSON, JARONE PINHASSI, RUI BRANCA, JANNE LEHTIÖ, ÅKE HAGSTRÖM 2016. Ecological significance of transporters in the Baltic Sea microbial foodweb. *In manuscript*.

MCCARTY, P. L. 2007. Thermodynamic electron equivalents model for bacterial yield prediction: modifications and comparative evaluations. *Biotechnol Bioeng,* 97**,** 377-88.

MEIER, H. E. M., ANDERSSON, H. C., ARHEIMER, B., BLENCKNER, T., CHUBARENKO, B., DONNELLY, C., EILOLA, K., GUSTAFSSON, B. G., HANSSON, A., HAVENHAND, J., HÖGLUND, A., KUZNETSOV, I., MACKENZIE, B. R., MÜLLER-KARULIS, B., NEUMANN, T., NIIRANEN, S., PIWOWARCZYK, J., RAUDSEPP, U., RECKERMANN, M., RUOHO-AIROLA, T., SAVCHUK, O. P., SCHENK, F., SCHIMANKE, S., VÄLI, G., WESLAWSKI, J.-M. & ZORITA, E. 2012. Comparing reconstructed past variations and future projections of the Baltic Sea ecosystem—first results from multi-model ensemble simulations. *Environmental Research Letters,* 7**,** 034005.

MIDDELBOE, M., HAGSTRÖM, A., BLACKBURN, N., SINN, B., FISCHER, U., BORCH, N. H., PINHASSI, J., SIMU, K. & LORENZ, M. G. 2001. Effects of Bacteriophages on the Population Dynamics of Four Strains of Pelagic Marine Bacteria. *Microb Ecol,* 42**,** 395-406.

SAVCHUK OLEG P. , B. G. G. A. B. R. M. L.-K. 2012. BALTSEM - a model for decision support within the Baltic Sea Region. Baltic nest Institute.

THINGSTAD, F. 2000. Elements of a teory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnology and Oceanography,* 45**,** 1320-1328.