Grant agreement number: call2012-88 (BLUEPRINT)
Project acronym: BLUEPRINT
Project title: Biological lenses using gene prints

Periodic report 2014

Period covered: 1.1.2014 – 31.12.2014

Project coordinator: Prof. Lasse Riemann, University of Copenhagen, Denmark
Phone: +45 35321959 (office)
Email: lriemann@bio.ku.dk
Project website: http://blueprint-project.org/

Preface

The BLUEPRINT (Biological lenses using gene prints) project started in January 2014 and will be running for four years. BLUEPRINT is funded by the BONUS programme through the European Community's Seventh Framework Programme (FP/2007-2013) under implementation agreement R&I/I3/2012/BONUS made with BONUS, the joint Baltic Sea research and development programme.

Year 1 of the BLUEPRINT project was busy for all project members, with the kickoff meeting in January, hiring of personnel, various field activities including the cruise activities, various conferences, workshops, and outreach activities. We have made strong progress in all work packages, including the generation and evaluation of environmental data, partial establishment of the bioinformatics pipeline, and partial implementation of generated field data into biogeochemical modeling. The results presented in this report are solid fundaments for the ongoing activities of BLUEPRINT, which will allow for a systematic evaluation of combined genomics and biogeochemistry to describe the environmental status of the Baltic Sea pelagial.

February 2015
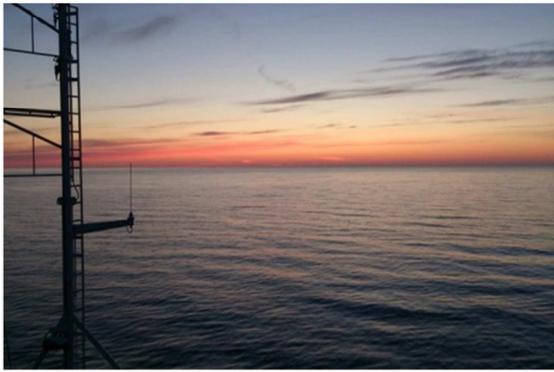*Lasse Riemann (project coordinator)*
Deniz Bombar (secretary).

Table of contents

1. Some highlights from year 1

2. Background and Objectives

Prokaryotic microbes are principal drivers of carbon and nutrient biogeochemistry and account for a major fraction of pelagic biomass and productivity in the Baltic Sea. Still, these organisms are neither included among the indicators of environmental status currently in use nor considered as functional entities in biogeochemical models. This flaw has been highlighted by HELCOM and OSPAR in their work to coordinate the development of indicators and determining good environmental status (GES) in the Baltic and North Sea areas. The last decade has witnessed a tremendous increase in the capacity of high-throughput technologies for retrieving and processing genetic information from environmental samples; this has given insights into microbially driven food-web processes and how microbes are affected by environmental conditions. Therefore, for the first time we now have the potential to make integrated use of this analysis capacity in a cost-efficient manner for developing a conceptual and methodological framework for the assessment of ecological status of the Baltic Sea ecosystem based on genetic information related to microbial functions and processes. Thus, this project will combine field studies, experiments, next-generation sequencing, bioinformatics and modeling to achieve the overarching objective: to establish a capacity to reliably deduce Baltic Sea environmental status based on indicators reflecting the biodiversity and genetic functional profiles of microbes in seawater samples.

3. Management and Organization

3.1 Management of BLUEPRINT

The BLUEPRINT project is organized as an interdisciplinary and multi-national consortium, which combines researchers with exceptional expertise in microbial oceanography and extensive experience with studying the Baltic Sea pelagic ecosystem. BLUEPRINT researchers merge the competence, geographical distribution around the Baltic Sea, and infrastructure needed to meet the complex challenges associated with the research, output coordination, and political implementation. We have implemented co-supervision of student and PostDoc projects in order to promote cross-disciplinary communication. Continuous involvement of stakeholders and end-users allows for underway project adjustments and efficient knowledge transfer from the project. Progress and internal communication is ensured by the coordination staff (UCPH) through travel, Emails, Skype meeting, meetings, and the project website (http://blueprint-project.org/). Scientific and organizational decisions are made by the Management Board. The independent international Advisory Board will continue to provide input on project progress, management, dissemination. For further details on project management in year 1 please see section 10.

3.2 Staff affiliated with the BLUEPRINT project

University of Copenhagen (UCPH), Denmark

Lasse Riemann, project coordinator, work package 7 leader

Mathias Middelboe

Trine Markussen, postdoc

Ryan W. Paerl, postdoc

Deniz Bombar, postdoc (not BLUEPRINT)

KTH Royal Institute of Technology (KTH), Sweden

Anders Andersson, work package 4 leader

Johannes Alneberg, PhD student

Luisa W Hugerth, PhD student (not BLUEPRINT)

Yue Hu, PhD student (not BLUEPRINT)

Stockholm University (SU), Sweden

Christoph Humborg

Bärbel Müller-Karulis, postdoc

Linnaeus University (LNU), Sweden

Jarone Pinhassi, work package 3 leader

Åke Hagström, work package 5 leader

John Larsson, postdoc

Daniel Lundin, postdoc (not BLUEPRINT)

Christofer Karlsson, PhD student

Therese Oscarsson, research assistant (not BLUEPRINT)

University of Helsinki (UH), Finland

Kaarina Sivonen

Jonna Teikari, PhD student

Hao Wang, postdoc

Leibniz Institute for Baltic Sea Research Warnemünde (IOW), Germany

Klaus Jürgens, work package 1 leader

Matthias Labrenz, work package 6 leader

Sara Beier, postdoc

Christin Bennke, postdoc

Alexander Hentzsch, research assistant

University of Tartu (UT), Estonia

Veljo Kisand, work package 2 leader

4. Scientific and/or technological results achieved during the reporting period

4.1 WP1: Spatiotemporal variations in the microbial BLUEPRINT of the Baltic Sea

Responsible partner: Klaus Jürgens, Leibniz Institute for Baltic Sea Research (IOW). Email: klaus.juergens@io-warnemuende.de

The goal of activities in WP1 is to analyze the transcriptional response of Baltic microbial communities along temporal and spatial gradients in order to link certain environmental parameter with the functional profile of prokaryotes and specific microbial activities. The major environmental gradients present in Baltic Sea environments are vertical oxygen depletion and horizontal changes in the salinity.

The compilation of previously collected field samples which are suitable for DNA/RNA analysis is still in progress and will be included in deliverable D.1.2 which is due in month 24.

During 2014 our major focus was on performing experimental work and field-sampling campaigns. Depending on the sampling context, e.g. the spatial or temporal frame, 16S rRNA amplicon data, metagenome data or metatranscriptome data may be the most appropriate method to display dynamics of the environmental status. Experimental work can give a good indication for which method most significantly described the environmental status at which context. For this reason, we have started analyzing samples

from new experiments before working with other samples that were (partly) simultaneously taken from the field.

The successful RV ALKOR cruise in June (3.-20.6.2014) served to observe the diversity and function of microbial communities across the large range of environmental conditions found in the Baltic Sea. The cruise covered the whole Baltic Sea with the different subsystems, and 31 stations had been sampled with the appropriate technologies (the new *in situ* sampling system "AFIS"), being the basis for non-biased metatranscriptomic studies (see 4.6). Sampling encompassed a range of GES and sub-GES conditions for which molecular microbial signatures will be examined. The report of WP6 describes the cruise in more detail, since it also has been used to develop standardized sampling procedures.

Preliminary results from this cruise served to optimize experimental design in WP2/WP3 by providing typical concentrations of well-known stressors such as DOC/DON and different inorganic nutrients. The collective metadata obtained during our cruise further aided colleagues in WP5 in their development of the biogeochemical model, by extending the coverage to a wider geographical/salinity region than previously (Gotland Sea surface layer).

In addition to the extensive field-based sampling, we have also performed two experiments (transplant experiment, mixing experiment) and carried out two further field-sampling campaigns (oxyclines at coastal sites, oxyclines in the Gotland deep).

For all experimental and field sampling work contextual parameters (e.g., nutrient chemistry) have been measured and filters for later nucleic acid extraction have been prepared. In case of the transplant experiment, RNA for metatranscriptomics and DNA for metagenomics has been extracted and is currently processed for Illumina sequencing. The protocol for nucleic acid processing was tested and established in close collaboration with WP6, and discussed with WP4. This protocol is available on the internal part of the BLUEPRINT webpage and is used by all partners.

4.1.1 Transplant experiment

Little is known about how locally adapted bacteria cope with changes in salinity and which ecological strategies they employ in response to such changes. We have performed a full-factorial transplant experiment with bacterial communities from three salinity regimes of the Baltic Sea (freshwater, brackish, marine), which were collected from a cruise with RV Alkor organized by WP6. Samples were reciprocally incubated under each other's environmental conditions (Fig. 1).
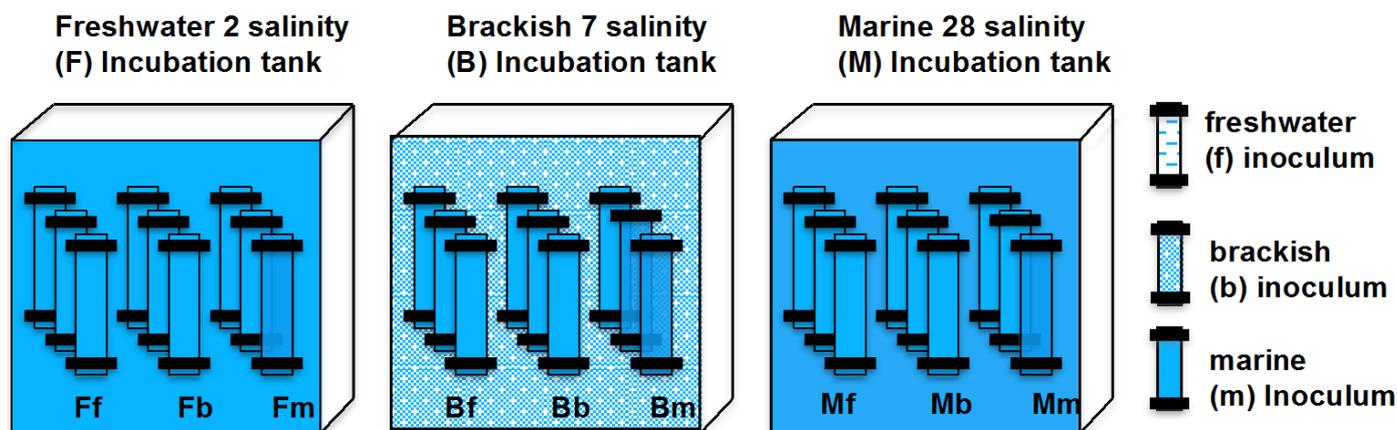

Figure 1: Experimental setup of the transplant experiment

We have daily measured cell numbers, overall metabolic activity via thymidine incorporation and the capacity for carbon substrate utilization (BIOLOG plates) to assess functional characteristics of the communities. Bacterioplankton was growing during the first three days of the experiment, then reaching stationary phase. Samples for nucleic acid extraction (for 16S rRNA and metatranscriptomic analysis) were taken at day 4 of the experiment.
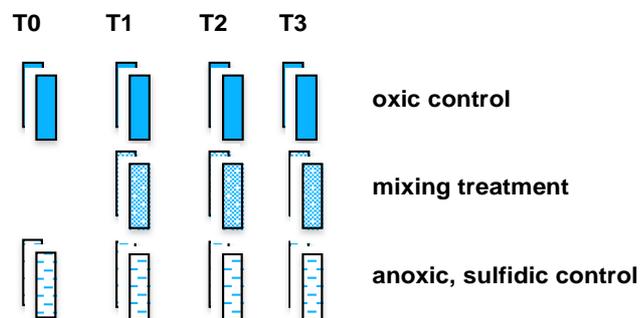
4.1.2 Mixing experiment

Whereas recent studies have analyzed metatranscriptomic data along profiles in marine oxygen minimum zones, there are no data available on the temporal transcriptional dynamics during mixing events between water layers with different redox characteristics. Small-scale mixing events are assumed to occur regularly in the redox transition zones (e.g., due to lateral intrusions or internal waves) and presumably strongly trigger microbial activity and important transformations, such as sulfide oxidation and denitrification.

Our aim is to use transcriptomic data in order to find characteristic gene expression patterns and transcriptional adaptions in response to changes (increase/decrease) in salinity. We further want to test for the importance of functional redundancy for community functioning at different disturbance and taxonomic levels, while hypothesizing that at increased levels of disturbance (salinity change) functional redundancy is shifted to higher taxonomic levels.

We have setup an experiment, where we have mixed adjacent oxic (O_2: 0.93 mg L^{-1}) and sulfidic layers (O_2: not detected; H_2S: 0.92 mg L^{-1}) of the oxygen minimum zone in the Gotland Deep (Fig. 2). These were sampled at three different time points after the mixing event (T1: ~1h; T2: ~4h; T3 ~9h).

During all time points (including T0), RNA-filters for metatranscriptomics were taken, as well as contextual data such as nutrient concentrations (O_2, H_2S, NO_2, NH_4^+, PO_4), prokaryotic cell numbers and bacterial production (via ^{3}H-leucine incorporation). Nutrient measurements indicated the gradual oxidation of sulfidic compounds in the mixing treatment, which is a

good precondition for our aim to study temporal dynamics and succession of bacterial mediated processes, induced by natural mixing events in the Baltic Sea.

Figure 2: Experimental setup of the mixing experiment

4.1.3 Field sampling campaigns

In order to assess consistencies and differences during oxygen depletion in the Baltic Sea at different conditions, and by comparing coastal and offshore sites, we have sampled oxygen depleted water bodies along oxyclines in coastal areas of the Baltics Sea (Boknis Eck, 2 profiles from sampling events in August and September) and in the central Baltic Gotland deep (4 profiles from different sites taken in October). For the samples contextual data as dissolved nutrients (NO_3^- NO_2^-, NH_4^+, SiO_2, PO_4^{3-}), cell numbers and different measured rates (nitrification in the coastal samples and bacterial production and bicarbonate incorporation in the Gotland deep samples) are available. Metatranscriptomic and environmental data from these samples will be available also for modeling in WP5

4.1.4 Outlook 2015

Our focus in 2015 will be to process and analyze metatranscriptomic data, which are currently being obtained in collaboration with WP 4, of the two performed experimental studies. We further consider accomplishing sampling from the Boknis Eck, by taking more samples during coastal anoxia, since in 2014 only on a single sampling day complete anoxia was reached.

4.2 WP2: Effects of environmental change on the microbial genetic BLUEPRINT

Responsible partner: Veljo Kisand, University of Tartu, Estonia. Email: kisand@ut.ee

This WP aims to experimentally identify the key genetic and functional properties of natural microbial assemblages exposed to defined stressors (e.g., hypoxia, N limitation leading to N2-fixing cyanobacterial blooms, and increased input of various sources of DOM) impacting the environmental status of the Baltic Sea. Increased input of DOM is predicted due to climate change, which causes increased precipitation in the catchment area, leading to higher river discharge. Cyanobacterial blooms, especially of toxin producing species, and hypoxia events are of increasing environmental concern in the Baltic Sea. How the above mentioned factors affect the fundamental biogeochemical processes modulated by microbes is not well studied. This WP focuses on experimental examination of the links between environmental conditions and BLUEPRINTs identified during in situ cruises. For comparative experiments we selected the most prominent stressors expected to expand in the Baltic Sea: (i) increased load of terrigenous DOM; (ii) cyanobacterial blooms; (iii) decreasing oxygen levels.

In 2014, two pilot batch experiments were carried out in Finland and Denmark in order to test and optimize method protocols as a preparation for full scale experiments in 2015 (more replication, multiple stressors, and sampling for molecular work). The full scale experiments in 10L containers will be performed at two sites using natural bacterial communities from the Gulf of Finland and the Øresund, Denmark. Moreover, a mesocosm experiment at LNU, Kalmar is scheduled in 300L containers in Fall 2015.

4.2.1 General set up of microcosms -Test experiment I
The first pilot experiment was carried out from 4.08 – 13.08.2014 at the Tvärminne Zoological Station,

Finland, with participation of UH and UT (Fig. 3). We tested the assembly of microcosms and setting up realistic experiments as a trial for experiments involving various experimental treatments,higher replication, and sampling for the genetic Blueprints. Natural seawater samples were collected in Storfjärden (Gulf of Finland) from the upper 3 m with the research vessel Lota. Nine different batches were set up and incubated at 21°C. Sampling occurred daily on five consecutive days to measure bacterial abundance, pH, concentrations of O2, nutrients and chlorophyll a, and to obtain samples for DNA/RNA extractions.

Figure 3: Upper panel: Sampling the Storfjärden near Tvärminne Zoological station for microcosm experiments in August 2014. Lower panel: Microcosms in the climate room.
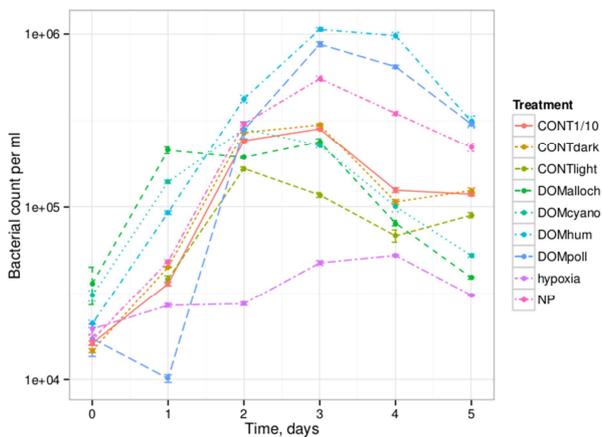
Figure 4: Time course showing the total number of bacteria in experimental microcosms. CONT1/10 - control with lower dilution ratio (1:10), CONTdark - control in the dark, CONTlight – control in the light, DOMauto – autochthonous dissolved organic matter, DOMhum – freshwater from humic rich river, DOMpoll – fresh water from polluted river, hypoxia – reduced O_2 concentration, NP – nutrient addition.

Setting up 10L batches proved suitable for the small scale experiments. As shown in Fig. 4, an experimental duration of 5 days was sufficient to cover the main responses of the bacterial community to the respective environmental stressors under summer conditions. Genetic samples served the purpose of optimizing DNA/RNA extraction protocols for subsequent use in the 2015 experiments. In future experiments we need to consider that additions of riverine DOM (rich in humic substances) also constitute a significant nutrient supply (e.g. ammonium), making it difficult to identify microbial responses solely caused by the actual DOM addition.

4.2.2 Measuring rates - Test experiment II

In December 2014, we carried out a test experiment (at UCPH in Helsingør) based on results and experiences from the Tvärminne summer experiment. The full scale (including proper replication and all treatments) Øresund batch experiment is planned to take place at the same location in April 2015. In December, 80 liters of surface water were collected off the coast of Helsingør and immediately brought back to the shore-based laboratory for preparation of the individual treatments (Fig. 5). The experiment was run at 12 °C for 6 days. The aim of the experiment was to test and optimize sampling and filtration procedures, as well as individual protocols not tested in the first experiment (e.g. bacterial production, exoenzyme activity, O2 respiration, and DOM bioavailabilty). Six batch cultures of 10 L were prepared, and samples taken daily in technical duplicates (Fig. 5).
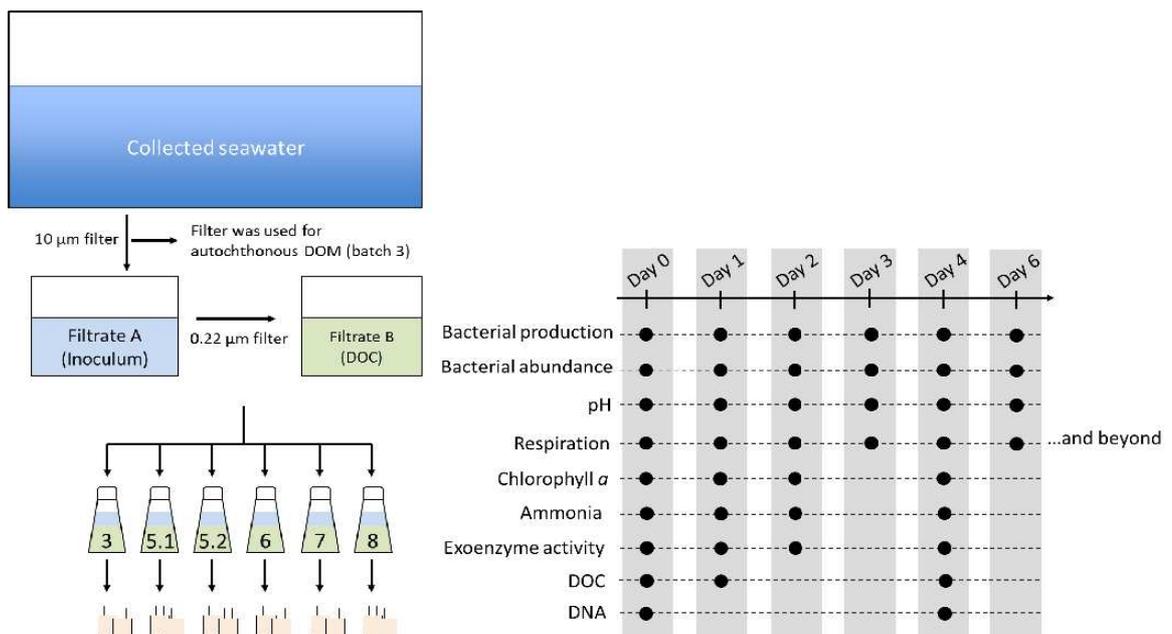
Figure 5: Treatments applied (left), 3 - autochthonous DOM in dark; 5.1 - Nutrient addition in dark; 5.2 - Nutrient addition in light; 6 - Reduced O_2 flow in dark; 7 - Control in dark; 8 - Control in light, and parameters obtained during daily sampling (right).

The observed response time of the bacterial community was even shorter than in the summer experiment, and bacterial production in the natural seawater used in this low productivity season appeared clearly carbon limited. Bacterial production was clearly different only in one treatment (Fig. 6), Three exoenzyme activities were measured: β-glucosidase activity; Alkaline phosphatase activity; and Protease activity. The activity levels of the three exoenzymes were clearly different between treatments (data not shown). The level of protease activity was the highest, while β-glucosidase activity was the lowest. Overall, the added DOM seemed to stimulate the activities of all three enzymes. Ammonium concentration increased during the experiment in all batches.
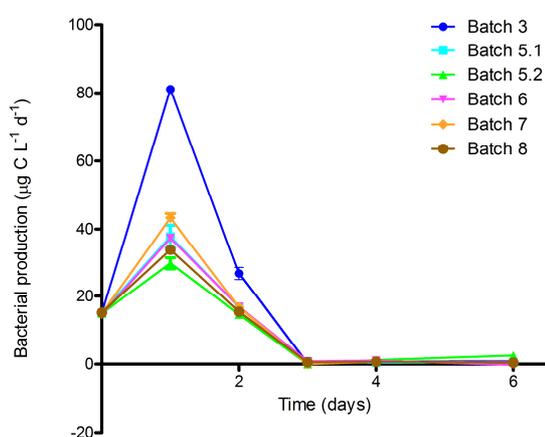
Figure 6: Bacterial production in Test Experiment II over the 6 day incubation period.

Adjustments for the full scale 2015 experiments include the revision of the incubation time for measuring β-glucosidase and alkaline phosphatase activity. In addition to the practical experience gained from these test experiments, we obtained useful information about appropriate incubation times, inoculum preparation, sampling frequency, and the preparation procedure of DOM additions. Moreover, they allowed us to evaluate the protocols for the specific measurements and adjust them to the abiotic conditions in Øresund.

4.2.3 Outlook 2015
Based on the pilot experiments, we are currently preparing the logistics and protocols for the full scale 10L experiments taking place in Helsingør (April 2015) and Tvärminne (August 2015). These experiments will apply identical experimental setups, with participants from UT, UH and UCPH. Moreover, these experiments also serve as preparation for the large-scale mesocosm experiment scheduled for September 2015 at LNU in Kalmar. It should, though, be mentioned that we are considering that postponing the mesocosm experiment to Spring 2016 would allow for better insights into the 2015 data prior to the mesocosm setup. That could potentially be invaluable information.

4.3 WP3: Impact of environmental stressors on the BLUEPRINT of model bacteria
Responsible partner: Jarone Pinhassi, Linnaeus University, Sweden. Email: jarone.pinhassi@lnu.se

4.3.1 Activities in 2014
The last decade, the development of novel sequencing technologies has allowed obtaining fundamental knowledge of genes and metabolisms important for aquatic bacterial communities in relation to their surrounding environment. Thus, it is now possible to study key genes that are involved in the utilization of different organic and inorganic nutrients, providing insight into the regulation of biogeochemical cycles.

Moreover, these key genes can be used as genetic signatures (i.e. marker/indicator genes) to define the functional ecology of aquatic bacterial communities and the environmental status of pelagic waters.

The objective of WP3 is to experimentally investigate responses in key diagnostic genes to selected external drivers/stressors. Responses of potential gene markers ("key genes") to external forcing mimicking environmental stressors on the Baltic Sea will be analyzed in experiments with well-studied Baltic Sea

model bacteria and cyanobacteria.

During the first year, a literature survey has been made to summarize current research that has produced insights into signature genes for biogeochemical processes, of potential importance for interpreting the responses of bacterioplankton to environmental forcing. The report resulted in a total of 225 genes, out of which 54 genes were nominated as having particular potential as indicator/marker genes. In the various WPs of BLUEPRINT, these genes will be given particular attention in the analysis of natural and experimental samples from the Baltic Sea.

Initial experiments have been made to determine bacterial responses to organic matter from phytoplankton species that form major blooms in the Baltic Sea, such as the cyanobacterium *Nodularia spumigena* and the diatom *Skeletonema marinoi*. Analysis of responses of natural Baltic Sea bacterial communities represent test beds for detailed analysis of specific bacterial model species. Initial results indicated that organic matter from the diatom constitutes a more favorable resource than from the cyanobacteria, and that phytoplankton derived organic matter substantially impacted bacterial community composition (Fig. 7). Thus, experiments to analyze gene expression responses of three selected Baltic Sea model bacteria to this organic matter were started.
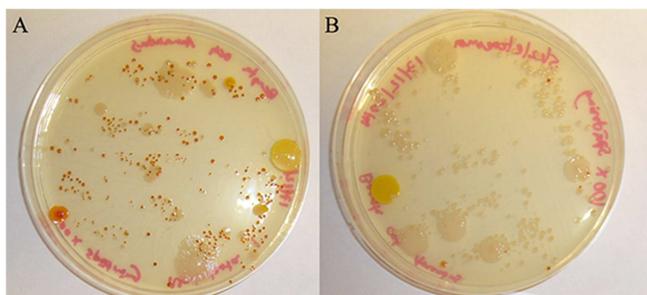
Figure 7: Bacterial colonies growing on agar plates when given organic matter input from: A) *Nodularia spumigena*, and B) *Skeletonema marinoi*. Note appearance of colonies with distinct morphologies and color, indicating divergence in composition of cultured species.

Regarding work on model cyanobacteria that are key species in the pelagic Baltic Sea ecosystem, work on

the major toxin-producing cyanobacterium *Anabaena* and the hepatotoxic *Nodularia* was performed. The main growth-limiting factor for these heterocystous autotrophs is inorganic phosphorus. During the year, studies on the effects of inorganic phosphorus depletion on the transcriptomes and proteomes of *Anabaena* sp. 90 (Fig. 8) have been carried out. This has been done to broaden the basic understanding of the importance of phosphorus to nitrogen-fixing cyanobacteria as well as to identify new marker genes for environmental monitoring. Exposure to prolonged phosphorus stress triggered the response of 823 differentially expressed genes and 43 differentially expressed proteins. A majority of these genes/proteins are involved in central metabolism and cellular growth. A large 1869 amino acid hypothetical protein was the most strongly up-regulated protein under phosphorus depletion warranting further study. In addition, 15 putative intergenic regions were differentially expressed, supporting previous studies of their importance as regulatory factors.
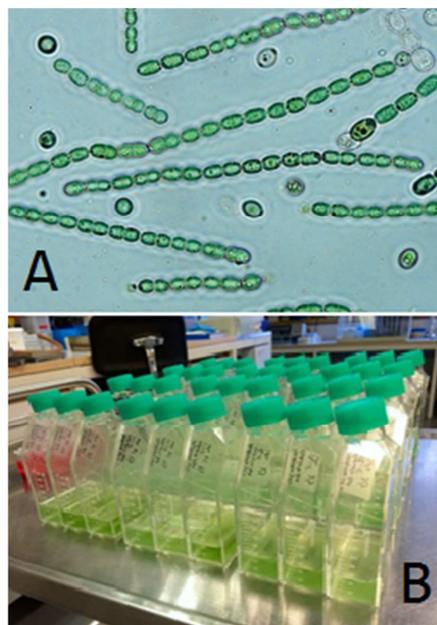
Figure 8: A) Images of *Anabaena* sp. 90 cells. B) experimental bottles with *Anabaena* sp. 315 growing in co-cultures with *Nodularia spumigena* sp. AV1.

4.3.2 Outlook 2015

During the coming year, detailed experiments with heterotrophic Baltic Sea model bacteria will examine their growth performance (biomass production and respiration; the latter will be assessed using a Picarro $\delta^{13}C$ carbon dioxide analyzer), in parallel with analysis of gene expression patterns, under different growth conditions. Also for some, environmental regulation of N_2 fixation will be examined. Experiments will involve exposure of bacteria to organic matter from cyanobacteria and diatoms, and to defined organic

macromolecules containing N and/or P. Oncoming experiments with Baltic Sea cyanobacteria will examine P-replete versus -deplete growth conditions, along with exposure to various salinities. Cyanobacterial growth will be examined along with gene expression analyses, to decipher marker genes for physiological stress, N_2 fixation and akinete production.

4.4 WP4: The bioinformatics platform

Responsible partner: Anders Andersson, KTH Royal Institute of Technology, Sweden. Email: anders.andersson@scilifelab.se

4.4.1 Brief description of WP4

The BLUEPRINT project aims at predicting the environmental status of a sample based on the taxonomic and functional properties of its microbial community. Work package 4 is responsible for building bioinformatic solutions for converting the metagenomic- and metatranscriptomic sequences into suitable data structures for downstream analysis, for storing this data in a database, and for predicting the environmental status of a sample from the data.

4.4.2 The BLUEPRINT Bioinformatics solutions

The bioinformatics workflow within the BLUEPRINT project can be divided into three main blocks (Fig. 9).

The first block concerns the generation of a metagenome reference assembly that should cover the major geochemical guilds of the Baltic Sea, based on metagenome sequencing data derived from sampling cruises and seasonal samplings at fixed stations. The short shotgun DNA sequences are collectively assembled into partial (hundreds to hundreds of thousands base-pair long) genomes. Subsequently genes are identified on these partial genomes, and are functionally and taxonomically

annotated, by comparing them with sequences in public databases. Since assembly is a computationally expensive process, the reference assembly will only be updated when significant amounts of new data are available.

The second block concerns how to obtain the "BLUEPRINT" of a sample, i.e. its functional and taxonomic profile. The BLUEPRINT can be based on either metagenome or metatranscriptome data. The sequence reads are "mapped" onto the reference metagenome and counts per gene are calculated. Combining this with the annotations of the genes, different types of taxonomic and functional profiles are calculated.

The third block concerns prediction of environmental status from the BLUEPRINT of a sample. Here an algorithm will be trained to either classify a sample in discrete classes (e.g. good-, intermediate- or bad environmental status) or predict a continuous value (e.g. the rate of a biogeochemical process) based on many training examples, i.e. BLUEPRINTs of samples with known classes/values.

During the first year of the project we have developed bioinformatic solutions for the first two blocks. A pilot Baltic Sea reference metagenome was constructed based on data from 33 seasonal surface water samples from the Linnaeus Microbial Observatory (LMO), 10 km east of Öland, collected during 2012. A bioinformatic pipeline was developed for extracting functional BLUEPRINTs from these samples. Further, a structure of a database for storing and retrieving BLUEPRINTs has been outlined. The work has been carried out in close collaboration with WP5.

4.4.3 Outlook 2015

We will expand the pipeline to also produce taxonomic BLUEPRINTs for the samples. We will also start developing a database with a user-friendly interface for storing and retrieving BLUEPRINT data.

Metagenome- and transcriptome sequencing will be conducted on samples obtained in a sample cruise of WP6 (June 2015). These samples are taken at different depths at stations covering the whole salinity gradient of the Baltic Sea. The metagenome data from these samples will be assembled into a BARM v1. Strategies for integrating the BLUEPRINT data into modelling will be tested together with WP5. Moreover, metagenome- and transcriptome sequencing will be done on samples from the large mesocosm experiment within WP2 scheduled for 2015. This will contribute to the establishment of links between genetic BLUEPRINTs and environmental condition/perturbation. In the analysis, particular attention will be devoted to the key indicator/marker genes identified in WP3 (Deliverable 3.1).
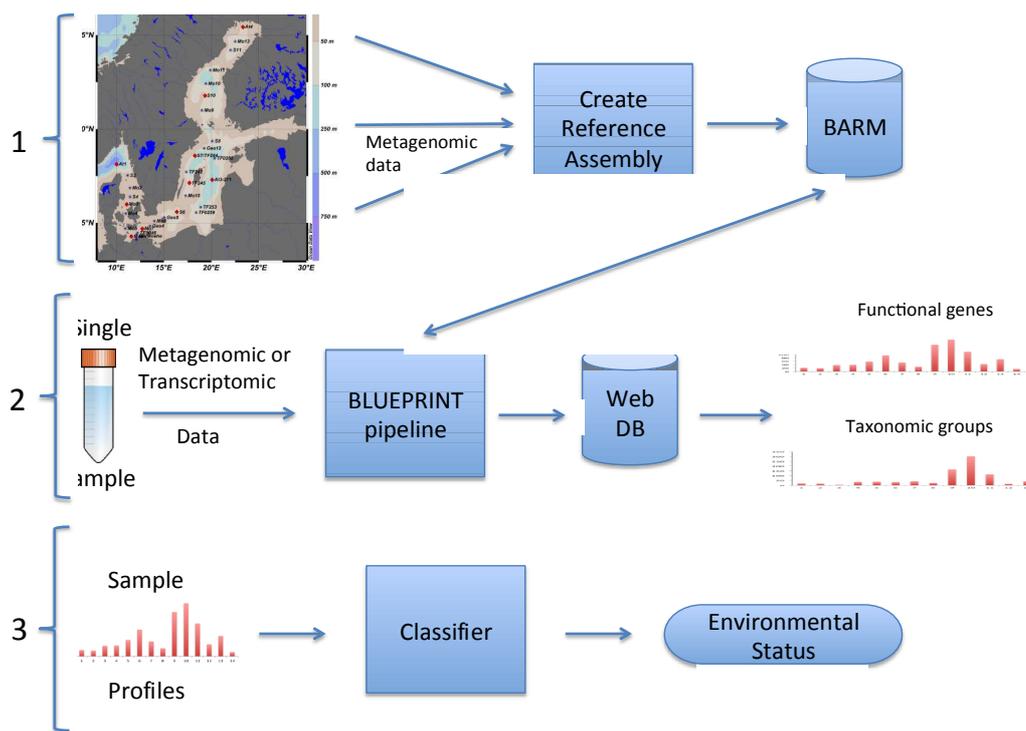
Figure 9: Outline of the bioinformatics workflow in the BLUEPRINT project. The numbers correspond to the three major bioinformatic blocks: 1) Assembly and annotation of a Baltic Sea reference metagenome (BARM); 2) Obtaining functional and taxonomic profiles ("BLUEPRINTs") from samples; 3) Predicting the environmental status of a sample from its BLUEPRINT.

4.5 WP5: Incorporation of BLUEPRINT in biogeochemical modeling

Responsible partner: Åke Hagström, Linnaeus University, Sweden. Email: ake.hagstrom@lnu.se

4.5.1 General work strategy

During the first year of BLUEPRINT work package 5, work has been focused on dissecting the flow of matter in the BALTSEM model in terms of its corresponding metabolic pathways. Two Baltic microbial metagenome datasets, one generated within the BLUEPRINT project (sampled in 2012) and one published in an earlier project (sampled in 2009; Dupont et al. 2014) have been used as references. In parallel the ability of the model to harmonize with specific data, generated within the project, has been explored in the form of "Model time-slices".

4.5.2 The Baltsem Model outline

The biogeochemical module of the BALTSEM model has recently been expanded to include dissolved organic carbon (DOC), nitrogen (DON) and phosphorus (DOP) as well as an explicit representation of H_2S and dissolved inorganic carbon (BALTSEM-C, Gustafsson et al. 2014). Three functional types of phytoplankton, one of them capable to fix atmospheric nitrogen, act as primary producers.

Heterotroph organisms simulate a community of particle feeders that use particulate organic matter for growth. Heterotrophs assimilate carbon, nitrogen and phosphorus according to a fixed elemental ratio.

4.5.3 Available model time slices

Simulated model time-slices were initially matched to observations at the Linnaeus microbial observatory (LMO), sampled by Linnaeus University, Kalmar. The station is located in the Western Gotland Sea 12 km off the coast of Öland (N 56°55.851, E 17°03.640).

The results showed that observations were occasionally close to the minima or maxima in the simulation range (Fig. 10). The loads used to force the model (present loads according to HELCOM PLC 5.5) were slightly lower than the calibration loads of the BALTSEM model for the same time-period. In addition, the simulations are close to equilibrium conditions between nutrient inputs and simulated nutrient pools whereas observations correspond to a transient

They excrete surplus nutrients to maintain homeostatic growth, and respire a temperature-dependent fraction of their food consuming oxygen.

Nutrients and carbon contained in dead autotroph and heterotroph organisms are assigned to a detritus pool. Release of DOC, DON and DOP from detritus followed by mineralization of the dissolved organic nutrients and carbon provides a secondary pelagic nutrient regeneration pathway. Depending on redox conditions either oxygen, nitrate (denitrification) or sulfate (sulfate reduction) is used as electron acceptor when dissolved organic matter is mineralized. In addition to these heterotroph processes, the model also simulates sulfate oxidation and nitrification rates to mimic chemoautotroph processes and a second pelagic denitrification pathway that couples nitrate reduction to sulfide oxidation.

Detritus provides the main sinking flux in the model and therefore links the pelagic and sediment compartments. In the present work we have restricted the analysis to the free water mass, thus the link to the bottom sediment will remain a potential capability.

eutrophication period. Furthermore, the observations correspond to a single station within a fairly large area that is treated as horizontally homogeneous in the model. Therefore, to include a wider range of nutrient concentrations into the model simulations BALTSEM was also forced with nutrient scenarios that represent plausible scenarios of increased and reduced nutrient loads compared to the present state. In addition, the BLUEPRINT June 2014 cruise (WP6) covered a transect through the entire Baltic Sea with ten stations selected for metagenomics analysis. The same model analysis on surface waters and the chemocline for these samples were run. So far model runs for BLUEPRINT were generated for an extensive combination of meteorological forcing and nutrient

loads to ensure that the available field data is within the simulation range. Simulation time for a 300-year run is approximately 2.5 hours; therefore additional model runs can be generated when necessary.

4.5.4 Linking mathematical modeling and genetics

WP5 initially intended to compare and integrate catabolic/anabolic pathways constructed from metagenomic and metatranscriptomic data with the flows of matter in the mathematical model structure. While this is still the main goal, we have also investigated different aspects of the first metagenome datasets in an attempt to find the optimal interface between model and genetic data. One such data interface is the functional genes encoding transporters, which can be seen as the start- and endpoints of different metabolic pathways.

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 a C:N:P ratio, the comparison between model and biochemistry pathway map is not as straightforward as first indicated.

Since the model depicts transfer of matter between organisms/groups while most metabolic maps show transformation of matter through biochemical processes within the cell, we need 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.

In the metagenomic data, the enzymes responsible for the transformation of matter in the metabolic pathways are recorded both in abundance and in a phylogenetic context. However, a large amount of other proteins, in particular membrane associated proteins such as transporters, are also recorded. In microbial membranes these proteins may constitute about 75% by weight and often make up the majority of expressed proteins in natural communities. The ongoing analysis of metagenomes from the LMO 2012 time-series have demonstrated that in addition to a clear annual succession that characterizes the prevalence of metabolic pathways in the microbial communities an equally strong temporal signal can be detected in the distribution of different types of transporters in the metagenomes obtained from the LMO 2012 time-series (Fig. 11).
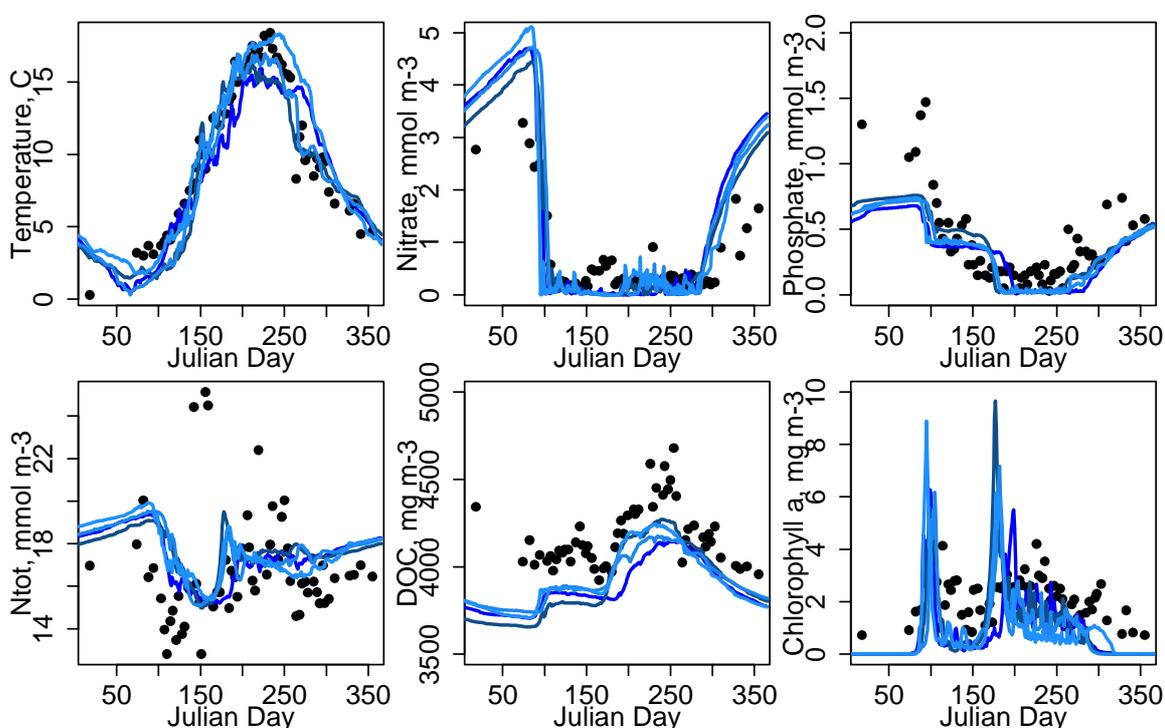
Figure 10: Best matching model time-slices for the 2012 seasonal dynamics at the LMO station for temperature, nitrate, phosphate, chlorophyll a, total nitrogen and DOC.

The significance of this observation 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.

Currently, WP 5 is gathering and summarizing genetic information on transporters in microbial genomes and in the near future, WP:s 1,2 and 5 will compare the abundance of different transporter genes to gene expression profiles (metatranscriptomics) as well as substrate uptake rates measured during the 2012 sampling. This will likely allow us to pinpoint the most significant transporters involved in the cycling of nutrients between microbial cells in the water column. Similar analyses of transporter variability have previously been used to analyze community response to phytoplankton blooms (Georges et al. 2014) and the effect of nutrient status on the microbial community (Zeigler-Allen et al. 2012).

4.5.5 Outlook 2015

The coming year 2015 WP5 will be occupied in an iterative discussion on how to merge metagenome data and mathematical modeling (primarily involving WP:s 3, 4 and 5). Initial development of the Metabolic Pathway Indicator (MPI) will start later in the year depending on data from planed fieldwork and will involve all partners and WPs.
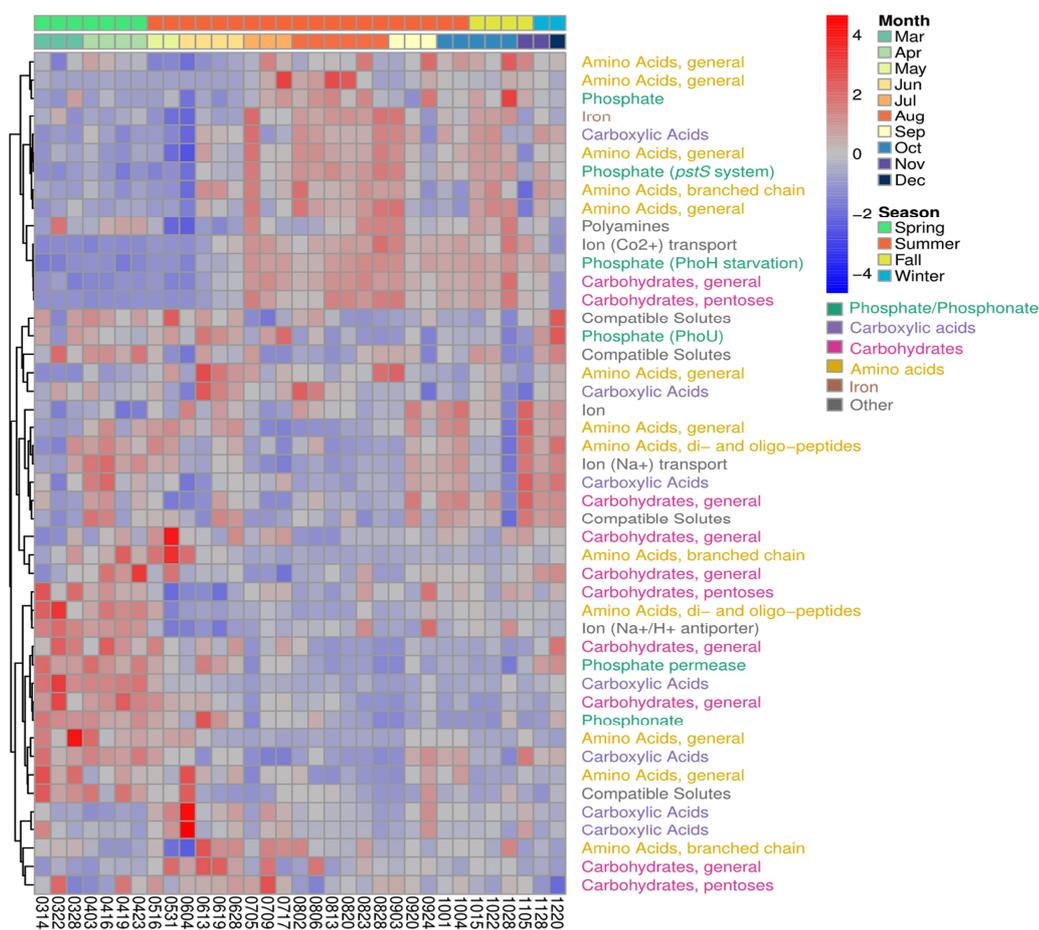
Figure 11: Heatmap showing the distribution of transporters in the metagenomes obtained from the LMO 2012 time-series. Here a subset of transporters, involved in uptake of various molecules such as carbohydrates, phosphate and amino acids, are shown. Each column represents a sample with collection dates (MMDD) shown at the bottom of the plot. Cell colours (blue = low to red = high) indicate relative abundance of each transporter across the season. Colored rows at the top indicate month of year and season (Spring - Winter). Rows are clustered using spearman rank correlations showing a clear seasonal succession in transporter abundance.

4.6 WP6: From mechanistic to functional monitoring – guiding microbial BLUEPRINT into practical operability

Responsible partner: Matthias Labrenz, Leibniz Institute for Baltic Sea Research (IOW). Email: matthias.labrenz@io-warnemuende.de

WP6 focuses on transfer and integration of new microbial descriptors into existing monitoring procedures. Several aspects have to be considered – going deep into very different disciplines – ranging from basic science over bioinformatics and potential development or adaptations of new instrumentations. WP6 will merge these aspects by improving, developing, evaluating, and standardizing general and specific workflows to guide BLUEPRINT into practical operability within the BLUEPRINT Competence Center (BCC).

4.6.1 Monitoring discussion forum

Scientists and experts in the field of monitoring met at the Leibniz Institute for Baltic Sea Research in Warnemünde (IOW), Germany, to discuss the approach of BLUEPRINT integrating microbial indicators into monitoring on December, 1st 2014. Several members of the discussion forum already participated at the kickoff meeting in January 2014. Additionally, national and international monitoring partners as well as engineers involved in the technical development of an *in-situ* sampler participated at the discussion forum. In total about 30 people from Germany, Denmark, Sweden, Poland and Estonia participated.

As an outcome of the 1st BLUEPRINT discussion forum it was proposed that BLUEPRINT will focus on:

1. investigative monitoring
2. identification of short and long term indicators
3. development of the Baltic reference database trying to define the GES
4. identification of suitable study areas with environmental gradients to train the algorithms
5. sampling at a fixed station (e.g. MARNET stations) to study short and long term changes.
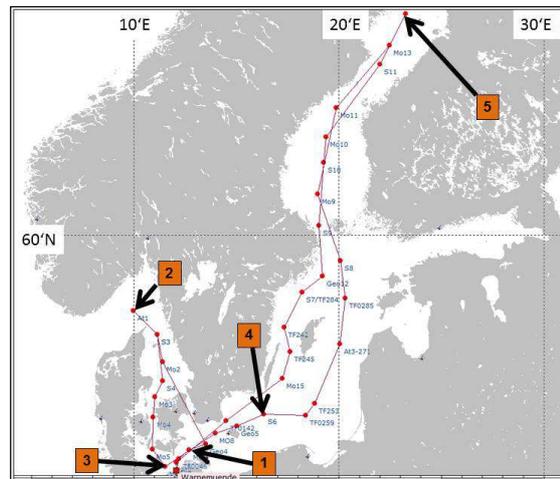
Figure 12: Station map of the BLUEPRINT cruise in June 2014. Highlighted stations represent the experimental stations with longer station time to set up specific experiments.

4.6.2 Standardization of BCC procedures

Standardized nucleic acid sampling, extraction, and sequencing (together with WP1 or WP4, respectively) procedures are fundamental for the BLUEPRINT monitoring approach. Therefore, a protocol/pipeline was established which serves as a template for the BLUEPRINT consortium and later for the BCC. The evaluation of the whole pipeline started in April 2014. Together with participants of WP7 and members of the BONUS Innovation project AFISmon, standardized water samples were taken during the RV ALKOR cruise in June (3.-20.6.2014, Figs. 12, 13) using the "automatic flow injection sampler" (AFIS, Figs. 14, 15) taking advantage of a fixative to preserve the status quo.

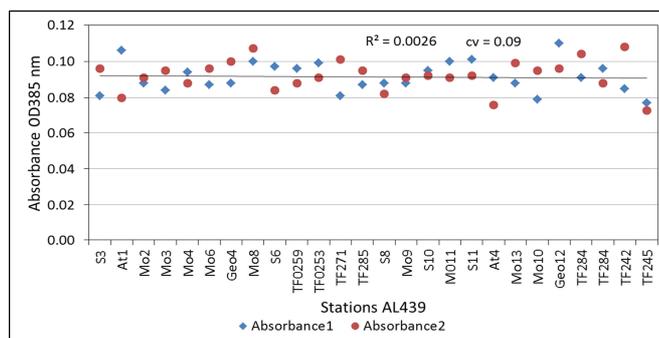**Figure 13:** RV Alkor 439 cruise participants.

Approval of the AFIS system for different areas of the Baltic Sea were done by adding ink as an internal standard to directly prove that injection of the fixative was successful. The ink concentration was measured photometrically in duplicates for each station (Fig. 15). The variation in its absorbance is negligible (cv = 0.09) and demonstrates the efficiency of the AFIS system in taking samples for metatranscriptome analysis.

Figure 15: Absorbance of ink concentration added to the AFIS-fixative at each station throughout the Baltic Sea transect.

Figure 14: On the left side the "old" AFIS sampler is depicted. This sampler was used during the AL439 cruise (left). The next graph shows the Prototype of the newly developed AFIS sampler, which will be used for upcoming sampling opportunities (middle). The graph on the right side represents the fixed Marnet station Darßer Schwelle (right).

Further on, at specific stations (deep oxic/anoxic zones, high salinities etc.) a comparison of non-fixed, board-fixed and in situ fixed samples was applied to prove the importance of sample preservation if meta-transcriptomic will be applied.

Following the standardization procedure the RNA was extracted from those samples, which are now queued for sequencing. With regard to metatranscriptomics and the decision of sequencing total RNA or mRNA, the first set of samples have been sent to the sequencing facility and are right now in progress.

4.6.3 Outlook 2015

The upcoming second year provides again the opportunity to go on a cruise for sampling. However, this cruise will cover other areas of the Baltic Sea including coastal regions in contrast to the June cruise AL439 (Figure 13). Further on, it is planned to sample at a fixed station (Marnet station) studying short and long term changes, implementing one of the outcomes of the first BLUEPRINT discussion forum into operability. For the standardized sampling during the 2nd cruise as well at the Marnet stations the newly developed AFIS sampler (Hydro-Bios) will be applied (Figure 14). This will result in a close collaboration with the project AFISmon, where an automatic water in situ fixation sampling prototype is currently developed for a first test run at the Marnet station.

The second BLUEPRINT discussion forum is planned in combination with the Baltic Sea Science Congress (BSSC) in Riga. In collaboration with WP7 and AFISmon a round table is planned at the BSSC discussing the indicator system for Baltic Sea monitoring. This gives us the opportunity to get in close contact with HELCOM discussing the approach of BLUEPRINT and the results of the first discussion forum.

4.7 WP7: Project management and political implementation

Responsible partner: Lasse Riemann, University of Copenhagen, Denmark. Email: lriemann@bio.ku.dk

We have strictly followed our chosen mechanisms to ensure effective internal communication, progress evaluation, and discussion of pertinent questions. The kickoff meeting in January 2014 was a great success in introducing all project partners, establishing the planning of close collaboration among WPs, and obtaining valuable first input from the independent international Advisory Board. Our project website includes an internal section only accessible to BLUEPRINT partners and has already proven as an effective communication organ for distributing method protocols, project guidelines, and dissemination products (http://blueprint-project.org/). In addition to the regular communication by email, the Skype meeting mid-2014 was a valuable opportunity to brief all project partners on the progress in the different WPs and to discuss *ad hoc* questions. The first discussion forum held in December 2014 at IOW, Germany, proved to be an extremely valuable exchange between BLUEPRINT scientists and monitoring experts. The upcoming 2. Year BLUEPRINT meeting is scheduled for 23-25 March in Stockholm. Key issues on the agenda are the continued integration of WPs and planning of activities involving multiple partners, like the scheduled large experiments. Moreover, the meeting will be a good opportunity for all the recently employed young scientists, postdocs and PhD students, to present their preliminary results and planned activities.

5. Amendments to the description of work and schedule of deliverables

At the moment we do not propose any major changes in the working plan, financial plan or the deliverables. However, we would like to mention difficulties with illumina sequencing that recently arose when project partners submitted samples from WP1 and WP6 to SciLife Lab in Stockholm. Unfortunately, a proportion of the samples failed in the sequence library preparation. All of these were subjected to removal of ribosomal RNA in order to sequence only mRNA, but the exact reasons for the failure of this step are currently unclear. This calls for very careful assessment of the next steps and optimization of the rRNA removal protocols, which may delay sequence data analysis and possibly dependent deliverables. We are working closely with the staff at SciLife and will do our best to overcome these difficulties as soon as possible, and will timely inform BONUS on any changes to the deliverables schedule, should they be unavoidable.

6. Promoting an effective science-policy interface to ensure optimal take up of research results (corresponding with the reported performance statistics 1-4)

6.1 Number of international, national and regional stakeholder events organized by the project (performance statistic 4)

1st Blueprint discussion forum, 1. December 2014, Institute for Baltic Sea Research Warnemuende (IOW)

The first Blueprint discussion held at IOW, Germany, aimed at bringing together Blueprint scientists, policymakers, legislators, stakeholders, and monitoring experts in order to evaluate the scientific content, practicability, and development of Blueprint, and also to provide support for monitoring and technical development (see 4.6.1). The event hosted 27 participants including stakeholders representing four different research institutions, three federal agencies, and a representative of industry. These discussion forum meetings are planned to happen once per year.

Further scheduled activities include presentations at the 10th Baltic Sea Science Congress in Riga, June 2015, as well as a round-table discussion on indicators. Moreover, we are responsible for a panel discussion at the 14th Symposium on Aquatic Microbial Ecology (Uppsala, Sweden, August 2015) on Baltic Sea ecology and the future of molecular microbial ecology. Finally, BLUEPRINT and its ideas will be presented at the ICES Annual Science Conference 2015, September 2015, Copenhagen. Activities at these meetings will ensure communication to and with a diverse array of stakeholders involved in marine monitoring and Baltic Sea management.

7. Collaboration with relevant research programs and the science communities in the other European sea basins and on international level (corresponding with the reported performance statistic 5)

Nothing to report at this stage

8. Progress in comparison with the original research and financial plan, and the schedule of deliverables

While the scientific progress is consistent with original research plan, only about 2/3 of the budget for Year 1 was spent. This was caused mainly by delayed employments of postdocs or PhD students and unanticipated problems with next-generations sequencing. This has generated delays in getting samples analyzed and consequently delayed use of funds. Beneficiaries have therefore transferred funds to subsequent years, within the respective categories, and no negative effects on the quality, quantity or timing of scientific outputs are anticipated.
Deliverables of Year 1 are listed below:

Deliverable 7.1: Launching of the BLUEPRINT webpage (Delivery month 3)

Deliverable 3.1: Compilation of signature genes indicative of GES/sub-GES conditions from literature on model microorganisms and natural seawater samples (Delivery month 6); available at http://blueprint-project.org/publications/eliverables/

Deliverable 4.1: Establishment of a bioinformatics toolbox for functional analysis of sequences from experimental work on isolates and natural microbial communities (Delivery month 12)

Deliverable 5.1: Compilation of available model time slices presented in a protocol showing the steady state values of individual processes and evaluation of sample range (Delivery month 12). Model time-slices can be extracted and analyzed online at (http://193.11.28.219:3838/BALTSEM_TS_finder).

9. Dissemination

Publications in press / published (entered in performance statistic 8)

Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Lahti L, Loman NJ, et al. 2014. Binning metagenomic contigs by coverage and composition. Nature Methods advance online publication.

Publications in prep / submitted

Beier S, Shen D, Juergens K. et al. Functional redundancy along changing salinity. Manuscript in preparation.

Bombar D, Bennke C, Labrenz M, Riemann L et al.: Assessment of microbially-mediated nitrogen gain and loss terms in Baltic Sea redoxclines using metatranscriptomics and stable isotope rate measurements. Manuscript in preparation.

Bombar D, Labrenz M, Sørensen SJ, Morberg S, Riemann L et al.: Metagenomic characterization of uncultivated heterotrophic N_2-fixing microorganisms (diazotrophs) isolated from the Baltic Sea Redoxcline using flow cytometry cell sorting. Manuscript in preparation.

Larsson J, Müller-Karulis B, Karlsson C, Andersson A, Hugerth L, Alneberg J, Pinhassi J, Hagström Å. Seasonal shifts in transporter gene abundance and composition in Baltic Sea microbial communities. Manuscript in preparation.

Larsson J, Müller-Karulis B, Karlsson C, Andersson A, Hugerth L, Alneberg J, Pinhassi J, Hagström Å. Mapping gene expression to reconstructed microbial genomes reveals key players in Baltic Sea nutrient transformations. Manuscript in preparation.

Lindh M, Sjöstedt J, Ekstam B, Casini M, Hugerth L, Lundin D, Andersson AF, Andersson A, Legrand C, Pinhassi J. Bimodal occupancy frequency distributions in marine bacterial communities – the importance of regional dynamics. Manuscript in preparation.

Lindh M, Figueroa D, Sjöstedt J, Lundin D, Andersson A, Legrand C, Pinhassi J. Transplant experiments uncover Baltic Sea basin-specific responses in bacterioplankton community composition and metabolic activities. Manuscript submitted to Frontiers in Microbiology.

Muthusamy S, Lundin D, Branca RM, Baltar F, González JM, Lehtiö J, Pinhassi J. Comparative analysis of exponential and stationary phase proteomes in three marine model bacteria (Alpha- and Gammaproteobacteria and Bacteroidetes). Manuscript in preparation.

Teikari J, Österholm J, Battchikova N, Kopf M, Wahlsten M, Aro EM, Hess WR, Sivonen K. Transcriptomics and proteomics profiling of Anabaena sp. strain 90 under inorganic phosphorus stress. Manuscript submitted to Applied and Environmental Microbiology.

Conference talks and posters

Lindh MV, Sjöstedt J, Ekstam B, Legrand C, Baltar F, Hugerth L, Lundin D, Nilsson E, Andersson AF, Pinhassi J. Ecological patterns within spatio-temporal fluctuations of the bacterioplankton consortium in the Baltic Sea. ASLO/AGU Ocean Sciences Meeting. Honolulu, Hawaii, 23-28 February, 2014. Oral presentation.

Pinhassi J, Palovaara J, Akram N, Baltar F, Forsberg J, Bunse C, Pedrós-Alió C, González M, González J. Regulation of proteorhodopsin phototrophy in the flavobacterium Dokdonia sp. MED134. ASLO/AGU Ocean Sciences Meeting. Honolulu, Hawaii, 23-28 February, 2014. Oral presentation.

Branca R, Zhu Y, Muthusamy S, Andersson AF, Pinhassi J, Lehtiö J. Sample fractionation by high resolution isoelectric focusing and database fragmentation by peptide pI allows efficient marine metaproteomics of seasonal dynamics in Baltic Sea microbiomes. 10th Siena Meeting. From genome to proteome: 20 years of proteomics. Siena, Italy, 31 Aug - 4 Sept, 2014. Oral presentation.

Lindh MV, Figueroa D, Sjöstedt J, Lundin D, Andersson AF, Legrand C, Pinhassi J. Baltic Sea transplant experiments uncover distinct water mass-dependent responses in bacterioplankton community composition and activities to changes in salinity and dissolved organic matter. Swedish Marine Sciences Conference, Umeå, Sweden, November 2014. Oral presentation.

Riemann L. N_2 fixation by heterotrophic diazotrophs in temperate estuarine waters. Microbes in the Baltic: Small things, small sea, big questions; 18-21 November 2014, Gdynia, Poland. Oral presentation.

Larsson J, Müller-Karulis B, Andersson AF, Hugerth L, Alneberg J, Pinhassi J, Hagström Å. Using metagenomics to find biological indicators and predict biochemical processes in the Baltic Sea. Microbes in the Baltic: Small things, small sea, big questions; 18-21 November 2014, Gdynia, Poland. Oral presentation.

Karlsson C. Exploring genetic responses of marine bacteria to environmental challenges in the Baltic Sea: Biological lenses using gene prints (BLUEPRINT). Poster presentation at Computational molecular analysis summer school, 29 Sept-3 Oct 2014, Wilhemshaven, Germany.

Hugerth LW, Alneberg J, Larsson J, Pinhassi J, Andersson AF. NGS of the Baltic Sea microbiome - from diversity patterns to genome-ecology links. Microbes in the Baltic: Small things, small sea, big questions; 18-21 November 2014, Gdynia, Poland

Hugerth LW, Alneberg J, Pinhassi P, Andersson AF. A metagenomic analysis of planktonic Actinobacteria and Bacteroidetes in the Baltic Proper. Microbes in the Baltic: Small things, small sea, big questions; 18-21 November 2014, Gdynia, Poland

Hu Y, Karlson B, Andersson AF. Pico - to mesoplankton distribution along the 2000 km salinity gradient of the Baltic Sea. Microbes in the Baltic: Small things, small sea, big questions; 18-21 November 2014, Gdynia, Poland

Alneberg J, Bjarnason BS, Hugerth LW, Larsson J, Pinhassi J, Schirmer M, Ijaz UZ, Loman N, Andersson AF, Quince C. Clustering Metagenomic Contigs using CONCOCT. Microbes in the Baltic: Small things, small sea, big questions; 18-21 November 2014, Gdynia, Poland

Kisand V. Microbial food webs and carbon cycle in the Baltic, any significant changes in a discourse during the last decades? Microbes in the Baltic: Small things, small sea, big questions; 18-21 November 2014, Gdynia, Poland.

Larsson J, Müller-Karulis B, Andersson AF, Hugerth L, Alneberg J, Pinhassi J, Hagström Å. Using metagenomics to find biological indicators and predict biochemical processes in the Baltic Sea. Microbes in the Baltic: Small things, small sea, big questions; 18-21 November 2014, Gdynia, Poland

Labrenz M. The BONUS projects BLUEPRINT and AFISmon. Presented at "Rendez-Vous de Concarneau: where Industry meets Science in marine Biotechnology" as well as the Biological Institute Helgoland colloquium of the Alfred-Wegener Institute. Concarneau, France, 9-10th October and Helgoland, Germany, 9th July 2014.

Outreach and Media output

Jarone Pinhassi: "Bakterier kan motverka klimatförändringarna": Interview in Swedish regional TV-news "Smålandsnytt" and Interview in Swedish regional newspaper "Barometern" on 19 August 2014

Lasse Riemann: Article in the regional newspaper "Nordsjælland". November 2013

Lasse Riemann: Interview and article on the BLUEPRINT project, its visions and potential impact with the regional newspaper "Helsingør Dagblad". November 2013

Lasse Riemann at Rotary club Helsingør. Molecular tools for monitoring environmental status in the Baltic Sea; the BLUEPRINT project. 16 June 2014

Participation in other BONUS projects

Participation in sampling of waters in the Roskilde Fjord, organized by the BONUS project COCOA: Therese Oscarsson (research assistant), 4 days in 2014

Presentation of BLUEPRINT by coordinator Lasse Riemann at the COCOA kick-off meeting 10 March 2014, Jyllinge, Denmark

Collaboration with the AFISmon project on the sampling at the Baltic Sea cruise, June 2014 (see 4.6)

13. Reference list

Dupont CL, Larsson J, Yooseph S, Ininbergs K, Goll J, et al. 2014. Functional Tradeoffs Underpin Salinity-Driven Divergence in Microbial Community Composition. PLoS ONE 9(2): e89549. doi:10.1371/journal.pone.0089549

Georges AA, El-Swais H, Craig SE, Li WKW, Walsh DA. 2014. Metaproteomic analysis of a winter to spring succession in coastal northwest Atlantic Ocean microbial plankton. I*SME Journal*, 8, 1301-1313. doi:10.1038/ismej.2013.234

Gustafsson E, Deutsch B, Gustafsson B G, Humborg C, Mörth CM. 2014. *Journal of Marine Systems*, *129*(C), 289–302. doi:10.1016/j.jmarsys.2013.07.005

HELCOM (2011). *Fifth Baltic Sea Pollution Load Compilation (PLC-5)*.

Zeigler-Allen L, Allen E, Badger J, McCrow J, Paulsen I, Elbourne L, Thiagarajan M, et al. 2012. Influence of nutrients and currents on the genomic composition of microbes across an upwelling mosaic. *ISME Journal*, 6, 1403–1414. doi:10.1038/ismej.2011.20

Annex 1: Blueprint's research staff sorted by age class

Age group	PhD students		Post-docs		Assistants, lecturers, instructors and eq		Associate professors and eq		Professors and eq	
	F	M	F	M	F	M	F	M	F	M
<= 24	0	0	0	0	0	0	0	0	0	0
25 - 49	3	2	3	5	2	1	0	6	0	0
50 - 64	0	0	0	0	0	0	0	0	1	2
> 65	0	0	0	0	0	0	0	0	0	1