Marine Pollution Bulletin. doi:10.1016/j.marpolbul.2015.01.008

Title: "Metabarcoding approach for the ballast water surveillance – an advantageous solution or an awkward challenge?"

Anastasija Zaiko[a,b*], Jose L. Martinez[c], Julia Schmidt-Petersen[d], Deni Ribicic[e], Aurelija Samuiloviene[a], Eva Garcia-Vazquez[c]
[a]Klaipeda University, H. Manto 84, 92294 Klaipeda, Lithuania
[b]Coastal and Freshwater Group, Cawthron Institute, 98 Halifax Street East, 7010 Nelson, New Zealand
[c]Department of Functional Biology, University of Oviedo. C/ Julian Claveria s/n. 33006-Oviedo, Spain
[d]Biology Department, Marine Biology Section, Ghent University, Krijgslaan, 281/S8, B-9000 Gent, Belgium
[e]Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology. Laboratory Center, Erling Skjalgsons gt. 1, N-7006 Trondheim, Norway

Emails:
Anastasija Zaiko: anastasija@corpi.ku.lt
Jose L. Martinez: secuenciacion@spi.uniovi.es
Julia Schmidt-Petersen: julia.schmidtpetersen@gmail.com
Deni Ribicic: deni.ribicic@ntnu.no
Aurelija Samuiloviene: aurelija@corpi.ku.lt
Eva Garcia-Vazquez: egv@uniovi.es
*Corresponding author

Abstract:

Transfer of organisms with ships' ballast water is recognized as a major pathway of non-indigenous species introduction and addressed in a few recent legislative initiatives. Among other they imply scientific and technical research and monitoring to be conducted in a efficient and reliable way. The recent development of DNA barcoding and metabarcoding technologies opens new opportunities for biodiversity and biosecurity surveillance. In the current study, the performance of metabarcoding approach was assessed in comparison to the conventional (visual) observations, during the *en route* experimental ballast water survey. Opportunities and limitations of the molecular method were identified from taxonomical datasets rendered by two molecular markers of different degree of universality - the universal cytochrome oxydase sub-unit I gene and a fragment of RuBisCO gene. The cost-efficacy and possible improvements of these methods are discussed for the further successful development and implementation of the approach in ballast water control and NIS surveillance.

Key words: invasive species, Ballast Water Management Convention, COI, RuBisCO, high-throughput sequencing, *en-route* observation

Introduction

Non-Indigenous Species (NIS) are recognized as one of the greatest threats to biodiversity worldwide (IUCN 2000). These are non-native species introduced from outside of their natural, past or present distributional range, deliberately or unintentionally by humans or other agents (Martin and Hines 2008). Part of them may spread in the recipient region and become invasive, acting as biological pollutants with adverse effects on biological diversity, ecosystem functioning and socio-economic values (Elliott 2003, Olenin et al. 2007). Records of new observations and established NIS have been increasing steadily in different marine ecosystems during the two last centuries and are still rising. In European marine ecosystems, on average two new NIS records occurred annually during the past decade (Olenin et al. 2013). Taking into account that global shipping activities have increased dramatically over the last decades, with >20% annual growth rate (Endersen et al. 2008), shipping is believed to be one of the most important pathways for species introductions (i.e., transport of organisms in ballast waters and/or in sediments of ballast tanks and biofouling) (Wonham et al. 2001, Leppakoski et al. 2002, Hewitt et al. 2009). It has been estimated that the major cargo vessels annually transport nearly 10 billion tons of ballast water (Gollasch et al. 2002), with thousands of stowaway organisms being transported every day (Carlton and Geller 1993, Gollasch et al. 2000b). This results in worldwide NIS exchange and growing risks of marine biological pollution (Elliott 2003).

The significance of NIS transfer is presently acknowledged by international organizations and is addressed in a number of recent legislative initiatives (e.g. EU Strategy on Invasive Alien Species, Marine Strategy Framework Directive). In particular, the importance of ballast water as a vector of species translocation is recognized through the Ballast Water Management (BWM) Convention, adopted in 2004 by the International Maritime Organization (IMO). The BWM Convention is aimed to prevent, minimize and ultimately eliminate the transfer of organisms via shipping, through the control and management of ships' ballast water and sediments (IMO 2004). IMO has formulated a number of special regulations in order to reduce the risk of organism transport. It also implies scientific and technical research and monitoring to be conducted by Parties of the BWM Convention. This includes observation, measurement, sampling, evaluation and analysis of the effectiveness of any management measure applied as well as analysis of any adverse impacts caused by such organisms and pathogens that have been transferred through ships' ballast water. On the other hand, the robust detection of NIS is crucial for implementing timely and cost-effective management measures including pre-border or early incursion control (Simberloff 2001, Darling and Blum 2007, Darling and Mahon 2011). Therefore there is a demand for rapid, standardized, reliable and cost-effective diagnostic tools that are able to identify and quantify the full range of NIS assemblages (King and Tamburri 2010, Olenin et al. 2011). It is expected that this demand will increase tremendously in the nearest future after updated marine monitoring programs are launched and BWM Convention enters into force.

The accuracy and resolution of NIS data needed for different policy measures varies. For instance, Regulation D-2 Ballast Water Performance Standard limits the acceptable concentration of all organisms in the discharge of ballast water, disregarding the species composition. Therefore, the species identification is not required for compliance control tests (IMO 2004). Uneven distribution and damage of organisms during the sampling procedure may however bias the test results. Particularly challenging is BWM Convention compliance verification for organisms of minimum dimension ≥50 μm, since less than 10 viable organisms of that size per cubic meter are acceptable by BWM Convention (Gollasch 2006, Gollasch et al. 2007). Therefore, intensive and extensive sampling is needed to ensure that violations of BWM Convention for discharges are detected by direct ballast water measurements (King and Tamburri 2010). Application of alternative techniques such as remote sensors, flow cytometry and

molecular methods are being increasingly discussed and addressed in experimental ballast water surveys (Gollasch et al. 2007, Harvey et al. 2009, Briski 2012, Ojaveer et al. 2014).

On the other hand, simple counts of observed organisms do not provide any additional information on biosecurity risks and are of minor scientific value for bioinvasion researchers. More detailed information (including taxonomic structure) is highly recommended for development of pressure indicators (e.g. rate of anthropogenic transport of NIS propagules), environmental status assessment, port baseline surveys, and species-specific risk assessments (Gollasch et al. 2007, Olenin et al. 2010, 2011, Ojaveer et al. 2014). This requires specific taxonomic expertise which is costly and extremely laborious, both in terms of representative sample collection and identification of dispersive life forms of NIS – eggs or larvae (Darling and Blum 2007, King and Tamburri 2010). It is believed that the application of rapidly developing molecular methods can substantially improve species identification capacities and aid NIS surveillance in the nearest future (Mountfort et al. 2012, Kelly et al. 2014, Wood et al. 2014).

Among molecular techniques, traditional DNA-based taxon identification approaches (e.g., PCR-based fingerprinting, quantitative PCR, Sanger DNA sequencing) can be efficient for detecting and identifying targeted NIS (Bott et al. 2010, Darling and Mahon 2011, Mountfort et al. 2012, Collins et al. 2013). However, these methods are often limited to a single species detection, and therefore not effective enough for biodiversity assessment. The recent development of the high-throughput DNA sequencing technology, also called Next Generation Sequencing (NGS), opened new opportunities for life sciences in general (Ansorge 2009) and demonstrated a great potential in marine biological and environmental studies in particular (Chariton et al. 2010). The major advance offered by this approach is the ability to operatively produce large numbers of comparatively low-cost sequences. This opens many different application opportunities, including metabarcoding studies: species detection and identification from bulk samples, using species-specific gene markers - the DNA barcodes (Hajibabaei et al. 2011, Andersen et al. 2012). DNA barcoding and NGS have already been recommended as a prospective tool for identifying NIS from environmental samples (Mountfort et al. 2012, Kelly et al. 2014, Ojaveer et al. 2014, Wood et al. 2014).

Here we address the applicability of metabarcoding methodology for the biosecurity surveillance, and particularly detection of organisms in ships' ballast waters. For this purpose the performance of metabarcoding approach was assessed in comparison to the conventional (visual) methodology, during the *en route* ballast waters survey onboard R/V "Polarstern". Opportunities and limitations of the molecular approach were identified from taxonomical datasets rendered by two molecular markers of different degree of universality - the universal cytochrome oxydase sub-unit I gene (COI) (Herbert et al. 2003) and a fragment of RuBisCO (RBC) gene, designed for diatom identification (Stoof-Leichsenring et al. 2012). The cost-efficacy and possible improvements are discussed for the further successful development and implementation of the approach in ballast water control and NIS surveillance.

Methods

The study was conducted during the ANT XXIX-1 EUROPA cruise onboard R/V "Polarstern", hosted by the Alfred Wegener Institute for Polar and Marine Research (Germany). The vessel left Bremerhaven port on October 28 2012, called at Las Palmas, Gran Canaria, on November 4 and left the next day, crossed the equator on November 14 and ended the cruise on November 27 in Cape Town, South Africa (Fig.1). For the purpose of the ballast water experimental study, the aft ballast tank (70 m^3) was filled with the North Sea water on October 28, out of Bremerhaven port. At the time of the ballast water upload, water temperature and salinity were 13.1^oC and 34 ppt, correspondingly.

Twenty samples of the ballast water were collected daily via the sounding pipe (20 mm diameter) starting from the 2nd day of the cruise, October 29 till November 17 (Fig. 1). Ballast water was extracted from approximately 1.5 m depth, through the build-in ballast pump (operational pressure up to 6 bar, loading capacity *ca.* 20 L/min), taking care to thoroughly flush the pipes before sampling. For each sample, 100 L of ballast water (measured with clean 10 L buckets) were concentrated by filtering through a plankton net (30 cm diameter, 55 μm mesh size) and instantly analyzed using a stereo microscope (60x magnification). The observed organisms were counted and identified to lowest possible taxonomic level. Only undamaged individuals without decay signatures were assessed, assuming them being viable or recently alive. Additional samples on Days 2, 11 and 21 of the cruise were collected and vacuum-filtered through sterile 0.12 μm Nuclepore™ membrane, which was thereafter preserved with 96% ethanol and stored at 4°C until the further land-based metabarcoding analysis of the settled material.

Simultaneously with the sample collection, environmental conditions in the ballast water were recorded, measuring salinity, temperature, pH, dissolved oxygen concentration with Ysi Professional Plus Multimeter. The sea surface temperatures were recorded automatically by the build-in onboard sensors.

Genetic and bioinformatics analyses

The precipitates from membrane filters were collected with sterile blades, then DNA was extracted from the filters using QIAamp DNA Mini Kit (Qiagen) and following the manufacturer extraction protocol. COI was amplified using the universal primers miniCOI (Meusnier et al. 2008) for PCR amplification. The fragment RBC gene was PCR amplified using a pair of primers designed for diatom identification (Stoof-Leichsenring et al. 2012). The comparison between the NGS results obtained from the two markers served to further explore the biases in ballast water biodiversity assessment due to primer specificity.

High-throughput sequencing was performed using the next generation sequencing platform Ion Personal Genome Machine System (PGM. Lifetechnologies) at Sequencing unit of the Oviedo University. For multiplexing purposes, the PCR products were labeled separately for each sample using short DNA sequences. Libraries were constructed using the Ion Plus Fragment Library Kit (Lifetechnologies) and templates were obtained using the Ion PGM™ Template OT2 200 Kit (Lifetechnologies). The templates were loaded in a 314 chip and sequenced using the Ion PGM Sequencing 200 Kit v2 (Lifetechnologies).

The yielded sequences where filtered by length (between 130 and 200 bp for COI and 80-130 for RBC gene) and quality (+20) and taxonomic classification (best hit, max e value = 0.001, min percent identity = 90.0) was assigned BLAST-aligning sequences against NCBI database using QIIME platform (Caporaso et al. 2010).

After initial inspection, the sequences of the organisms unlikely to be present alive in ballast water (e.g. vertebrates and non-aquatic species) were eliminated from the dataset. Most probably, sequences of those organisms were derived from body remains such as scales or feathers that can occur in marine water uploaded in the tank, but they will have no biological significance as NIS.

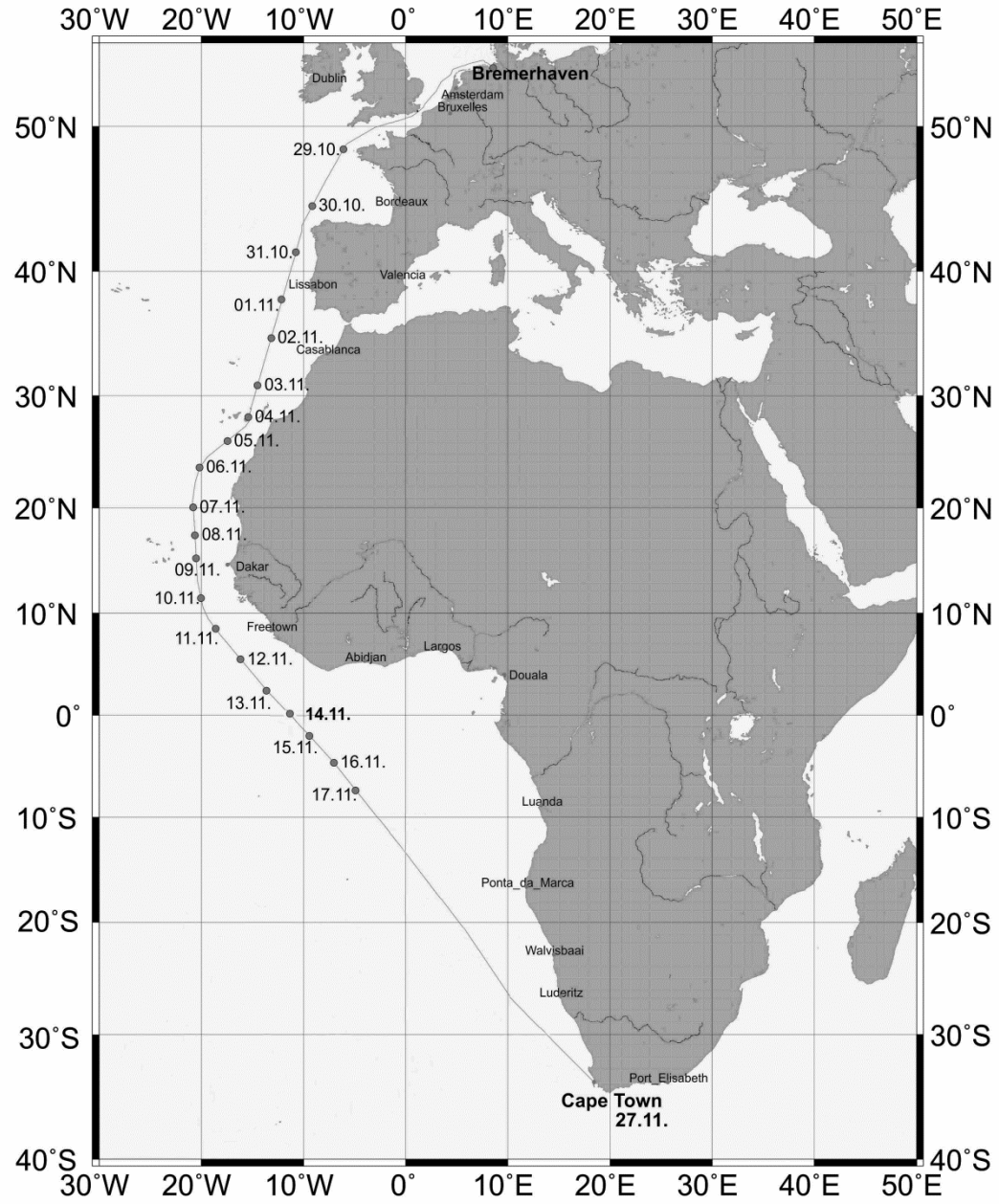

Figure 1. ANT XXIX-1 EUROPA cruise route with indicated days and locations of ballast water sampling.

Statistical analysis

For quantification summary and statistical analyses of biodiversity the species data were pooled to superior taxonomic ranks (Class or higher), as in Smith et al. (1999), for a conservative estimate of biodiversity in the ballast water assemblages. This way, possible uncertainties in the visual taxonomic assessment (due to for example the ambiguous phenotypes in species with phenotypic plasticity and uncertain identification of some larvae and algae propagules) were minimized. The composition of species (presence-absence data) identified from visual analysis was compared between sampling days using nonmetric multidimentional scaling (NMDS) based on Jaccard similarity matrix. NMDS cannot process identical samples, so a dummy variable (value 1) was added to all samples when constructing the similarity matrix. NMDS was

undertaken with 100 random restarts and visualized in two-dimensional plot. Linear regression was applied to ascertain the trends in densities of metazoans, protozoans and algae visually identified from ballast water samples.

To verify the differences in taxonomic diversity reported from visually analyzed samples and metabarcoding results (with COI and RBC barcodes), the canonical discriminant analysis of principal coordinates (CAP) was applied, followed by a permutation test and two-dimensional visualization. The Gower's distance measure was used for dissimilarity matrix construction, based on densities (number of observed individuals and number of yielded sequences for visual and metabarcoding data correspondingly) and standardized by total number per sample. The Gower's distance is considered robust enough for standardized data analysis, flexible and non-sensitive to missing observations and double zeros (Quinn and Keough 2002, Anderson and Willis 2003). Samples from days 16-19 with no visually detected specimens were excluded from the analysis.

The analyses were implemented in PRIMER 6 software package (PRIMER-E, Ltd., UK) and the R v3 statistical computing environment (R-project 2014).

Results

Environmental conditions within the ballast tank

The temperature within the ballast tank showed a steady increase over the first 14 days of observation, in consistence with the overboard temperature and reached the maximum of 29.9 ^{o}C on day 15. During the following six days the temperature dropped gradually by 6^{o} C, exceeding the sea surface temperature by 2^{o}C on average (Fig. 2). Dissolved oxygen concentration decreased from the normoxic conditions (7.3 mg/L; 91.5% saturation) at the beginning of the voyage, to anoxia (0.3 mg/L; 5% saturation) on the 20th day of observations. The pH values also demonstrated a gradual decrease in the range from 7.9 to 7.1.

Visual analysis

From the visually analyzed samples 14 taxa were identified during the observation period (Table 1). The noticeable change in community composition was reported starting from the second week of the observations (Fig.3).

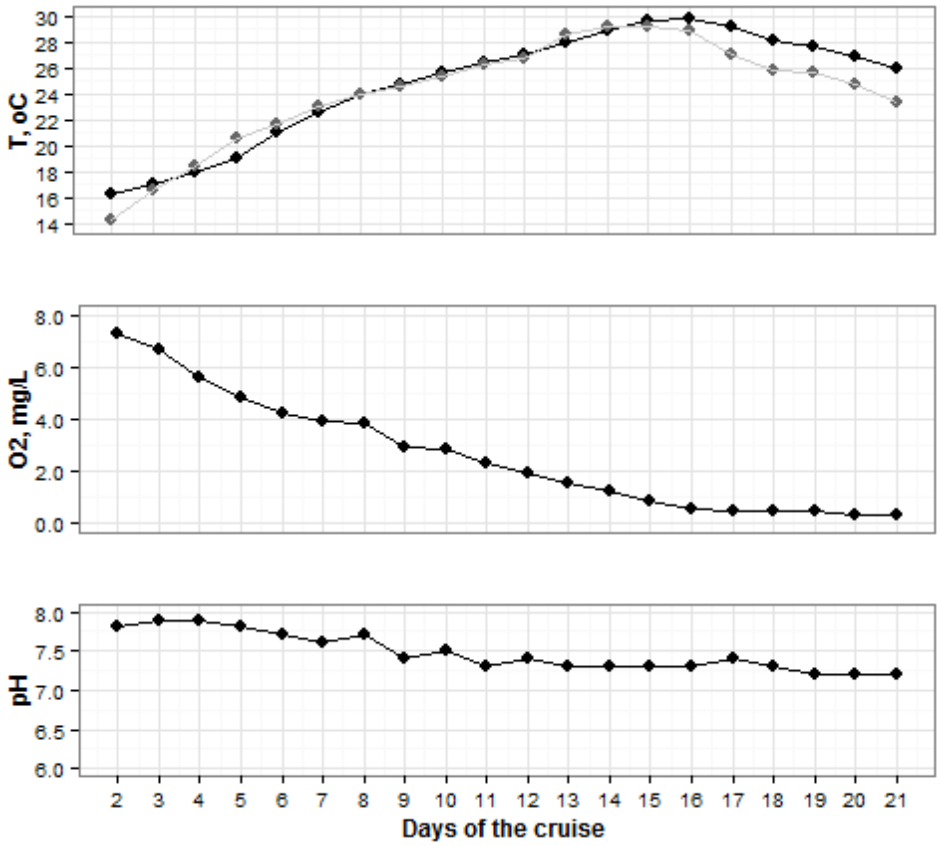

Figure 2. Registered environmental conditions in the ballast tank: water temperature (black dots -ballast water, grey dots – sea surface), dissolved oxygen concentration and pH.

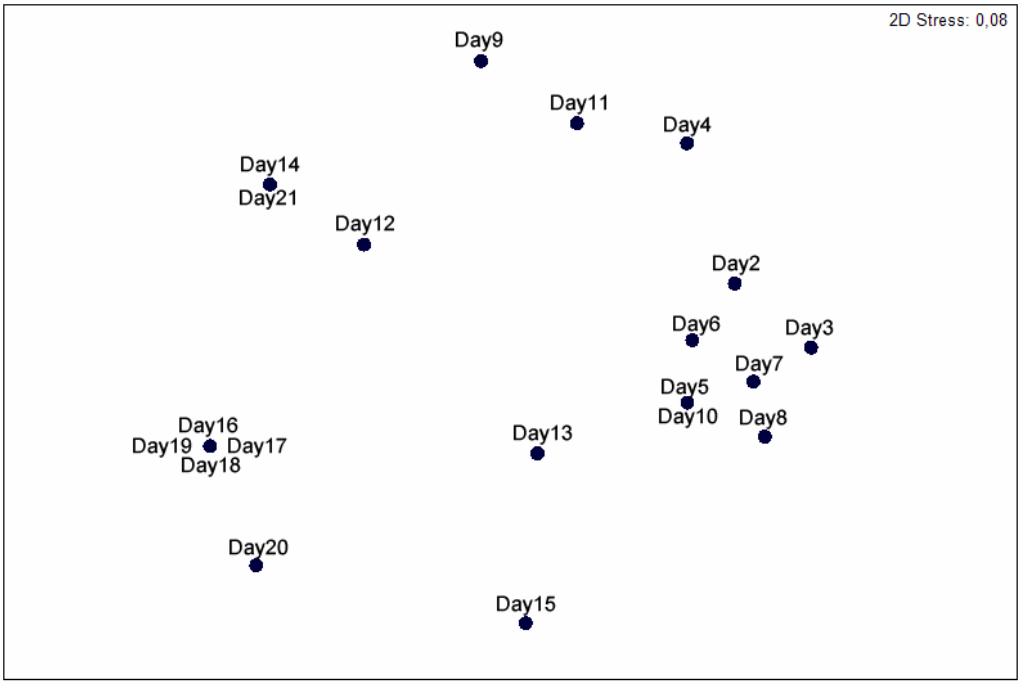

Figure 3.NMDS plot of the visually observed ballast water community, based on presence/absence of observed taxa (Jaccard index similarity matrix).

Table 1. Reported biodiversity from the ballast tank during the cruise (represented by the number of observed specimens and number of yielded sequences for the visual and metabarcoding data correspondingly, counts per 100 L of ballast water). COI1, COI2, COI3 and RBC1, RBC2, RBC3 samples correspond to Day 2, 11 and 21 samples. Unident. Invert: Unidentified invertebrates.

Taxa	Visual observations (days of voyage)																				Metabarcoding samples					
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	COI1	COI2	COI3	RBC1	RBC2	RBC3
Algae (total)	**457**	**309**	**107**	**82**	**17**	**37**	**15**	**10**	**5**	**6**	**0**	**2**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**113**	**15**	**45**	**325013**	**26476**	**33209**
Bacillariaphyceae	404	308	105	79	14	34	12	10	3	5	0	0	0	0	0	0	0	0	0	0	113	15	0	285687	22993	29056
Dinophyceae	53	1	1	3	3	3	3	0	2	1	0	2	0	0	0	0	0	0	0	0	0	0	0	12829	976	1044
Phaeophyta	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2306	110	118
Rhodophyta	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	45	86	8	63
Xantophyta	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	24135	2389	2928
Metazoa (total)	**17**	**4**	**2**	**4**	**5**	**7**	**6**	**1**	**2**	**1**	**2**	**5**	**0**	**9**	**0**	**0**	**0**	**0**	**5**	**0**	**12217**	**16068**	**2795**	**0**	**0**	**0**
Annelida	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0
Chaetognatha	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	5	0	0	0	0	0	0	0
Copepoda	13	0	0	3	3	1	1	0	1	0	2	5	0	7	0	0	0	0	0	0	0	0	17	0	0	0
Cnidaria	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Mollusca	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1154	179	0	0	0
Nematoda	1	2	2	0	2	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ostracoda	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Porifera	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1871	2564	474	0	0	0
Rotifera	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	412	0	0	0
Unident.invert.	2	1	0	1	0	3	1	0	1	0	0	0	0	1	0	0	0	0	0	0	10344	12344	1856	0	0	0
Protozoa	**1**	**0**	**1**	**0**	**0**	**0**	**0**	**1**	**0**	**1**	**1**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**122**	**49**	**11**	**862**	**0**	**0**	**0**
Oomycota	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**310**	**17**	**4611**	**0**	**0**	**0**

In the beginning, samples were dominated by numerous microalgae (diatoms and dinoflagellates), with a significant negative trend (R^2=42%, p=0.002) - more than tenfold decrease in a week and further drop down to zero values after few more days (Fig.4). The last record of algae (dinoflagellate) cells was reported on the 13th day of the cruise (9 November).

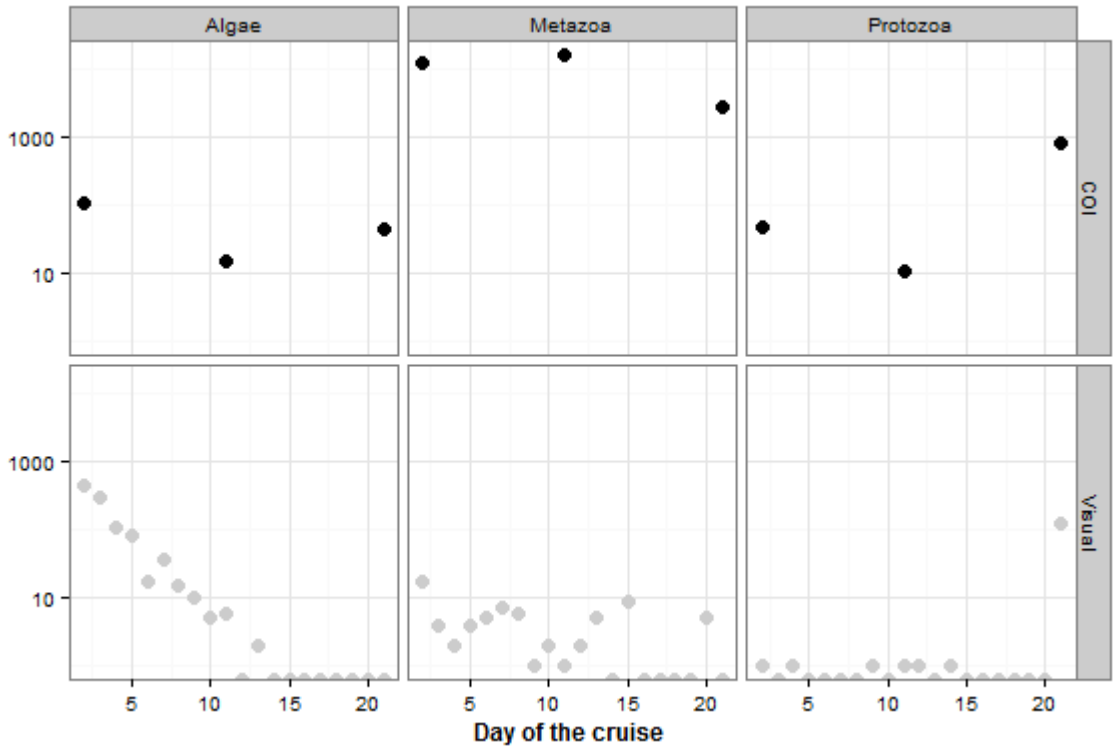

Figure 4. Densities of algae, metazoan and protozoan organisms reported from the visual analysis of the ballast water samples (number of individuals counted from the 100 L sample) and metabarcoding analysis with COI barcode gene applied (number of sequences yielded from NGS).

The metazoan taxa have also demonstrated significant negative trend (R^2=26%, p=0.02) over the observation period, yet individual organisms (arrow worms) were reported from the samples on day 20. It is worth noting however that increasing number of damaged, partly decayed individuals was registered starting from the second week of observations (day 8 and thereafter).

Single protozoan specimens were steadily reported from the samples with no apparent trend in densities (R^2=14%, p=0.09), except for an abrupt outbreak on the last sampling day when 122 actively moving protozoans with cilia-like protrusions were observed in the sample.

NGS results

PCR amplifications of both COI and RuBisCO (RBC) genes resulted in amplicons (145 and 97 nucleotide long correspondingly) from Day 2 (COI1, RBC1), Day 11 (COI2, RBC2) and Day 21 (COI3, RBC3) samples. The raw high-throughput sequencing of the amplicons produced 159 039 (COI) and 137 518 (RBC) reads. The stringent quality check and filtering parameters resulted in the removal of 28.7% (COI) and 32.3% (RBC) sequences. The number of high-quality sequences used for the further downstream analysis was 113 267 and 93 060

for COI and RBC samples correspondingly. The bioinformatics analysis resulted in the clustering of sequences into 29 and 136 OTUs, for COI and RBC respectively, yielding positive assignment hits against NCBI database with >90% sequence similarity and identified at genus level for conservative approach (Supp.1 and 2, summarized in Table 1).

Absolute majority of assigned sequences (96%) from COI1 sample matched with metazoan species, while most (55%) of COI3 sample sequences were assigned to oomycetes (fungus-like eukaryotic microorganisms), followed by unidentified marine invertebrates (22%) and protozoans (10%). The RBC samples were highly dominated (nearly 90%) by diatom sequences as expected from primers' specificity, followed by yellow-green algae.

Generally, there was apparent decrease in a number of yielded sequences from Day 21 sample for both applied markers (comparing to the Day 2 sample). However, when partitioned by large taxonomic groups based on COI results, a remarkable drop in number of sequences was apparent for algae (-60%) and metazoans (-77%), while protozoans demonstrated more than tenfold increase (Fig. 4). The oomycetes (not detected by visual analysis) have also noticeably increased in number of yielded sequences by the end of observation period.

At a lower taxonomic level, there were a few taxa that have demonstrated increase in sequence number by more than 100% over the observation period (Suppl. 1 and 2). The bigger increases of DNA sequences detected with COI primers were those assigned to the water mold *Achlya,* Arcellinidae protozoan *Hyalosphenia* and the rotifer *Brachionus* (Suppl. 1). The copepod *Cyclopodia,* red algae, protozoans, gastropods and most of the oomycetes have demonstrated somewhat increase in later samples. On the other hand, the algae diversity derived by NGS with RBC primers (Supp. 2) was more consistent among samples. Several diatom genera (namely, *Eunotogramma*, *Minidiscus*, *Skeletonema* and *Thalassiosira*) did show higher than ten-fold increase in the number of sequences during the observation period. Red algae and yellow-green *Botrydiopsis* have also demonstrated moderate increase by Day 21.

The diatoms detected with COI primers (*Nitzchia spp.*) were reported from the RBC samples as well. However red algae assignments showed discrepancy between two applied markers. Two genera of Ceramiales (*Polysiphonia* and *Dasya*) and *Plocamium* (Plocamiales) were obtained with COI, while RBC has resulted in *Ceramium* (Ceramiales) and *Delisea* (Bonnemaisoniales).

The permutation test carried out by CAP produced a p-value of 0.001 (based on 999 permutations). It means that no randomly permuted data set had more extreme assemblage than that of the original data set. It is noticeable that visually analyzed samples from Days 2-11 with a major component of algae were located on the positive-valued side of the Axis 1 (Fig. 5). Expectedly, the RBC samples with applied diatom-specific barcode grouped on the positive-valued side as well. The taxonomic diversity reported from the ballast water samples on 12th day of the cruise onwards grouped on the negative-valued side of the Axis 1, correlating with protozoan abundance and moderately – with metazoans and oomycetes. Yet, the COI samples were apparently distinct from others, demonstrating the highest positive correlation (ranging between 0.4 and 0.6) with the Axis 2 and correlating mostly with metazoans and oomycetes.

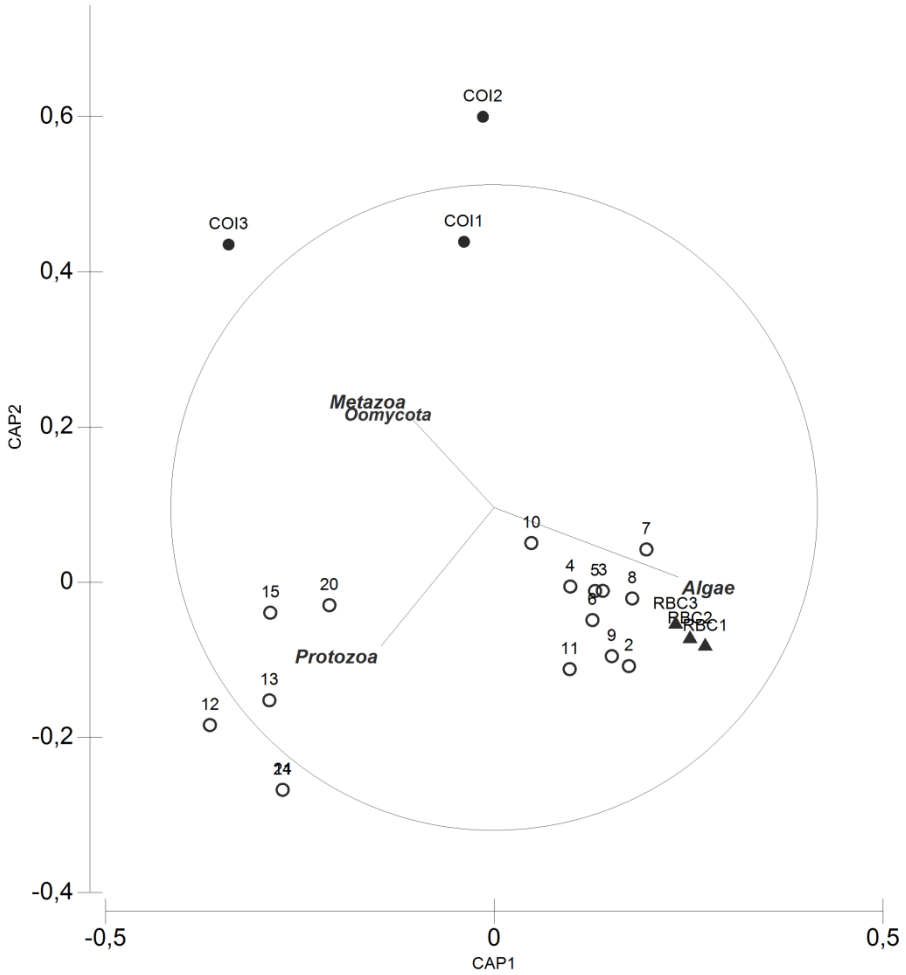

Figure 5. The first two canonical axes of the CAP analysis, based on taxonomic diversity identified from the ballast water samples applying visual analysis, COI and RBC barcodes (only major taxonomic groups are displayed for clear reading). The labels indicate the sampling day for visually analyzed samples (open circles) and sample code for RBC and COI samples (black triangles and filled circles correspondingly).

Discussion

This study evidences again that despite rather harsh environmental conditions (darkness, low oxygen, temperature fluctuations) during the long cross-latitudinal voyage there is a possibility for some eukaryote species to survive and even flourish in ballast waters (Gollasch et al. 2000a, 2000b, Olenin et al. 2000, Duggan et al. 2005). Such organisms are likely to remain viable upon discharge, thus posing a high risk of incursion to a recipient ecosystem. The combination of metabarcoding and conventional (visual) taxonomic analysis let us assess the dynamics in ballast water plankton community and identify taxa that have the highest potential of survival.

In general, the biodiversity revealed by both approaches at a higher taxonomic level (Table 1) coincided with that reported from other studies (Gollasch et al. 2000a, 2000b, 2002, Olenin et al. 2000, Duggan et al. 2005, Flagella et al. 2007, Briski et al. 2012). As reported also by other researchers, a rapid decline in plankton abundances was noticed during the first several days of the voyage, with a more pronounced decrease for the phytoplankton taxa (Gollasch et al. 2000a, 2000b, Olenin et al. 2000). The metazoan organisms have demonstrated higher persistence with

evident drop at the most adverse environmental conditions within the tank (anoxic, high temperature) during the second week of the cruise (Figs. 2, 3 and 4). Protozoans however were able to withstand the harsh ballast water environment and even increased in abundance by the end of the observation period. Such apparent shift in the ballast water community from algae-dominated to metazoan- and further microplankton-dominated (Fig. 5) can be explained by phytoplankton and zooplankton mortality increasing with time due to exposure to stressful conditions, and proliferation of saprophagous water molds and protists resistant to high temperatures, lack of light and oxygen (Jobard et al. 2010). On the other hand, these organisms were largely underestimated in previous ballast water surveys (Gollasch et al. 2000b, 2002, Duggan et al. 2005, Flagella et al. 2007) and in most NIS inventories in general (Wyatt and Carlton 2002). For instance, the absolute majority of species currently reported as associated with vessel vectors in the AquaNIS database are metazoans (Olenin et al. 2014). Small multicellular and unicellular organisms (as well as dispersible living stages of many taxa) are easy to overlook and confound in the conventional taxonomic assessment (Foissner 2006) or due to unrepresentative sampling strategy (Gollasch et al. 2007). In this case metabarcoding and NGS technologies are advantageous being able to detect and identify species from a single cell present in the sample (Jerde et al. 2011, Kelly et al. 2014).

At the lower taxonomic level, as resulted from the NGS data, there were a few genera that have not been reported from the ballast water surveys previously (Supp. 1 and 2). Some of those are known to contain NIS or even invasive alien species. For instance, the red algae *Dasya baillouviana* recorded in the Baltic Sea since 1960s (Maggs and Stegenga 1999), several species of *Polysiphonia* are recognized as NIS from North Atlantic, Mediterranean, Australia, New Zealand and Japan (Hurd et al. 2004, Minchin 2007, Geoffroy et al. 2012). However, the exact pathway or vector of spread is still largely undetermined for those species (Thomsen et al. 2007). The invasive oomycetes of *Phytophthora* genus are known to be associated with plant twig blight disease in Europe (Werres et al. 2001) and Sudden Oak Death disease in USA (Rizzo et al. 2005). Although it is believed that *Phytophthora* species are distributed predominantly via the terrestrial pathways (wind, land-based transport, planting material from infected nurseries), they are known to remain viable in water for years (Ko 2003) thus potentially could be transported with ballast water loaded from estuarine or coastal areas.

In the current study we did not assess specifically the viability of the organisms observed, as it is required for BWM Convention compliance control (e.g. Regulation D-2). In visual analysis we assumed that entire, undamaged individuals are likely to be alive at the sampling time or shortly before. However DNA molecules can resist for some days inside dead cells and even naked until degradation, as demonstrated in some previous studies of environmental DNA (e.g. Dejean et al. 2011). Therefore finding evidence of DNA from a species in environmental samples does not mean that it belong to a living organism. This, and inability of providing the measure of minimum dimension of observed organisms are probably the main weaknesses of metabarcoding application for the BWM Convention compliance control. However, DNA can be a signal of living organisms when its density increases with time. This could be the case of *Achlya*, *Brachionus* and *Hyalosphenia* in the present study (Supp. 1). On the other hand, an increase of these taxa would be not surprising, since *Achlya* belongs to hardy water molds (Willoughby 1965), *Brachionus* adults have been previously found surviving the long voyages within the ballast water (Gollasch et al 2002; Duggan et al. 2005) and *Hyalosphenia*, a widespread and rather resistant representative of testate amoebas (Heger et al. 2013), known to be transported in ballast water as far as to the Great Lakes (e.g. Nicholls and MacIsaac 2004). Among other

organisms that have not decreased (or even slightly increased) in density (Supp. 1 and 2) as evidenced by both by COI and RBC markers, the red algae (e.g. *Polysiphonia*) have free-living life-history phases (Kaczmarska and Dowe 1997) and can tolerate high temperatures, low salinities and lack of light for extended periods (Fralick and Mathieson 1975); diatoms are capable of survival in darkness 12 days and more (laboratory experiment at 18^{o}C temperature, Jochem 1999), maintaining cell abundance up to 90 days (laboratory experiment, at 15^{o}C temperature, Smayda and Mitchell-Innes 1974) and reported repeatedly from the ballast water samples (Gollasch et al. 2002).

However, there is another pitfall in metabarcoding application, related to the method's quantification capacity. Although eDNA concentration and number of sequences yielded from NGS are positively correlated with biomass or population density, estimates of absolute abundance remain elusive (Kelly et al. 2014). So far this approach cannot be applied independently for robust quantification and assessment of surviving taxa, but rather used as additional technique for biodiversity screening (e.g. if there are some doubts of non-compliance with BWM Convention, or a need for species-based risk assessment). For the putative samples or taxa more detailed further molecular analysis would be advised.

The use of multiple markers is often recommended for metabarcoding purposes, since it allows reducing amplification bias (Kelly et al. 2014). For instance, in the current dataset some inconsistences between visual analysis, RBC and COI could be explained by the specificity of the applied primers (e.g. Wilcox et al. 2013). In NGS results, copepods, arrow worms, nematodes were highly underrepresented. On the other hand, water molds, amoebas and rotifers were largely overlooked in the conventional analysis, but detected with COI marker instead. Genetic detection of most algae was possible only with the specific RBC primers. Indeed, the difference between the results obtained here from specific (RBC specific for diatoms; Stoof-Leichsenring et al. 2012) and more generalist (miniCOI; Meusnier et al. 2008) primers is enormous. Without the specialist primers diatoms, but also green and yellow algae, would remain inadvertent or highly underestimated in this study.

Since true universal primers annealing with same preference to all living taxa do not exist yet, we would suggest using primers cocktails for targeting a wider taxonomic spectrum as recommended by other authors as well (e.g. Valentini et al. 2009; Ivanova et al. 2011). Particularly, targeting diatoms more specifically is highly desirable on the short voyage legs, while community remains phytoplankton-dominated (Fig. 5). Diatoms are generally overlooked in conventional ballast water surveys due to light silicification of some taxa or resting stages present (Antia and Cheng 1970; McCarthy and Crowder 2000). We would also suggest replicating NGS from the same environmental sample employing different primer sets, specifically designed for the more problematic taxonomic groups (Jerde et al. 2011; Wilcox et al. 2013) and more particularly those expected to be found and survive in ballast water. This might improve the utility of this novel methodology for ballast water management issues.

This study was an experimental survey and not a real experience of ballast water monitoring. For the robust risk assessment, a bigger sampling effort, targeting different areas of the tanks and including sediments is recommended (Gollasch 2006). The results of a survey (counts of living biota) might also be affected by the uneven distribution of organisms within ballast tanks, sampling induced damage and mortality (particularly when pumping the water through the sounding pipe) or organism loss during the sample concentration (Gollasch et al. 2007). Hence, the application of more sensitive and specific of molecular techniques would be particularly advantageous as a complementary measure for species detection and identification. NGS

application has resulted in interesting findings and provided significant added value to the study outcome, even within the comparatively small-scale experiment. As a conclusion we would like to summarize the strengths and weaknesses of metabarcoding application for ballast water surveys in comparison to the conventional (visual) approach (Table 2).

Table 2. Preliminary comparison of strengths and weaknesses of metabarcoding and visual analysis approaches in ballast water surveys (based on the current study experience and literature review).

	Metabarcoding	Visual analysis
Cost	**Moderate:** $10,000-$50,000[1]/vessel/sampling voyage, with tendency to further decrease due to increasing NGS capacities	**Very high**: $75,000–$125,000/vessel/ sampling voyage[2]
Effort (time consumption)	**Moderate:** tendency to further decrease with rapid technology development[3]	**Low-Moderate:** depending on volumes and aims of the survey[2,4]
Quantification	**Moderate:** Increases with use of multiple markers, replication and internal control[5]	**High:** The accuracy however is highly dependent on sampling effort[4,6]
Taxonomic resolution	**High-Very high:** Increases with use of multiple markers, development of reference databases for barcode genes[5,7]	**Low-Moderate:** Highly dependent on taxonomical expertize of the researcher, low for early life stages (larvae, eggs, resting stages)[8,9]
Assessment of viability	**Low-Moderate:** Could be improved with sequential sampling (analysis of OTU dynamics) or RNA-based analyses	**Moderate:** Conservative assessment by visual inspection, intact individuals or dye-based methods (high uncertainty, particularly for certain groups and life stages)[6]
Sensitivity/detectability	**High-Very High:** Able to detect species at low abundances (individual cells, trace material)[10,11]	**Low-Moderate:** Rare or particularly small organisms are likely to be overlooked or underestimated[9]
Added value	**Very high:** Massive amount of "non-target" data on biodiversity might be obtained	**Low:** Generally, only targeted range of organisms (size- or taxonomically-based) is detected

[1]Costs are given as a proxy since may vary enormously depending on the particular molecular methodologies employed (price reduction is expected as long as new technical improvements), salary conditions, aim of the survey, vessel type, etc.; [2]King and Tamburri 2010; [3]Wood et al. 2014; [4]Gollasch 2006; [5]Kelly et al. 2014; [6]Gollasch et al. 2007; [7]Pochon et al. 2015; [8]Ardura et al. 2010; [9]Pochon et al. 2013; [10]Bott et al. 2010; [11]Darling and Mahon 2011.

Conclusions

Although based on a single experimental study, the results allow us to recommend some actions for improving the efficacy of ballast water surveillance and management employing next generation molecular technologies. First, using metabarcoding or any other molecular methodology for ascertaining the taxonomic status of organisms contained in ballast water samples is desirable, especially for taxa where microscopic identification is doubtful or very laborious. Second, ballast water monitoring is recommended to occasionally verify the positive results of control surveys (zero counts of ballast water organisms), because outbreaks of resistant species can be produced even in adverse conditions and overlooked by conventional analysis. Third, combining different methods (e.g. *de visu* surveys and metabarcoding) considerably improves the power of monitoring protocols.

Acknowledgments

We thank the crew of the RV Polarstern for their so valuable collaboration. This study has been supported mainly from the Campus of Excellence of the University of Oviedo, with additional funds from the Spanish National Project MINECO CGL2013-42415-R. This is a contribution from the Marine Observatory of Asturias. The contribution from the Klaipeda University was supported by DEVOTES (DEVelopment Of innovative Tools for understanding marine biodiversity and assessing Good Environmental Status) project funded by the European Union under the 7th Framework Programme, 'The Ocean of Tomorrow' Theme (Grant Agreement No. 308392), www.devotes-project.eu.

References

Andersen, K., K. L. Bird, M. Rasmussen, J. Haile, H. Breuning-Mdsen, K. H. Kjaer, L. Orlando, M. T. P. Gilbert, and E. Willerslev. 2012. Meta-barcoding of 'dirt' DNA from soil reflects vertebrate biodiversity. Molecular Ecology 21:1966-1979.

Anderson, M. J., and T. J. Willis. 2003. Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology. Ecology 84:511-525.

Ansorge, W. J. 2009. Next-generation DNA sequencing techniques. New Biotechnology 25:195-203.

Antia, N. J., and J. Y. Cheng. 1970. The survival of axenic cultures of marine planktonic algae from prolonged exposure to darkness at 20 C. Phycologia 9:179-183.

Ardura, A., A. R. Linde, J. C. Moreira, and E. Garcia-Vazquez. 2010. DNA barcoding for conservation and management of Amazonian commercial fish. Biological Conservation 143:1438-1443.

AquaNIS.Information system on Aquatic Non-Indigenous and Cryptogenic Species.www.corpi.ku.lt/databases/aquanis.Version 2.36+. Accessed 2014-10-28

Bott, N. J., K. M. Ophel-Keller, M. T. Sierp, Herdina, K. P. Rowling, A. C. McKay, M. G. K. Loo, J. E. Tanner, and M. R. Deveney. 2010. Toward routine, DNA-based detection methods for marine pests. Biotechnology Advances 28:706-714.

Briski, E., S. Ghabooli, S. A. Bailey, and H. J. MacIsaac. 2012. Invasion risk posed by macroinvertebrates transported in ships' ballast tanks. Biological Invasions 14:1843-1850.

Carlton, J. T., and J. B. Geller. 1993. Ecological roulette: the global transport of non-indigenous marine organisms. Science 261:78-82.

Chariton, A. A., L. N. Court, D. M. Hartley, M. J. Colloff, and C. M. Hardy. 2010. Ecological assessment of estuarine sediments by pyrosequencing eukaryotic ribosomal DNA. Frontiers in Ecology and the Environment 8:233-238.

Collins, R., K. Armstrong, A. Holyoake, and S. Keeling. 2013. Something in the water: biosecurity monitoring of ornamental fish imports using environmental DNA. Biological Invasions 15:1209-1215.

Caporaso, J.G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. Gonzalez Pena, J. K. Goodrich, J. I. Gordon, G. A. Huttley, S. T. Kelley, D. Knights, J. E. Koenig, R. E. Ley, C. A. Lozupone, D. McDonald, B. D. Muegge, M. Pirrung, J. Reeder, J. R. Sevinsky, P. J. Turnbaugh, W. A. Walters, J. Widmann, T. Yatsunenko, J. Zaneveld and R. Knight.2010. QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7: 335-336.

Darling, J. A., and M. J. Blum. 2007. DNA-based methods for monitoring invasive species: a review and prospectus. Biological Invasions 9:751-765.

Darling, J. A., and A. R. Mahon. 2011. From molecules to management: adopting DNA-based methods for monitoring biological invasions in aquatic environments. Environmental Research 111:978-988.

Dejean T, Valentini A, Duparc A, Pellier-Cuit S, Pompanon F, et al. (2011) Persistence of Environmental DNA in Freshwater Ecosystems. PLoS ONE 6(8): e23398. doi:10.1371/journal.pone.0023398

Duggan, I. C., C. D. A. van Overdijk, S. A. Bailey, P. T. Jenkins, H. Limen, and H. J. MacIsaac. 2005. Invertebrates associated with residual ballast water and sediments of cargo-carrying ships entering the Great Lakes. Canadian Journal of Fisheries and Aquatic Sciences 62:2463-2474

Elliott, M. 2003. Biological pollutants and biological pollution - an increasing cause for concern. Marine Pollution Bulletin 46:275-280.

Endersen, O., M. Eide, S. Dalsoren, I. S. saksen, and E. Sorgard. 2008. The Environmental Impacts of Increased International Maritime Shipping. Pages 43 *in* Global Forum on Transport and Environment in a Globalising World, Guadalajara, Mexico.

Flagella, M. M., M. Verlaque, A. Soria, and M. C. Buia. 2007. Macroalgal survival in ballast water tanks. Marine Pollution Bulletin 54:1395-1401.

Foissner, W. 2006. Biogeography and Dispersal of Micro-organisms: A Review Emphasizing Protists, ActaZoologica 45:111-136.

Fralick, R. A., and A. C. Mathieson. 1975. Physiological ecology of four *Polysiphonia* species (Rhodophyta, Ceramiales). Marine Biology 29:29-36.

Geoffroy, A., L. Le Gall, and C. Destombe. 2012. Cryptic introduction of the red alga Polysiphoniamorrowii Harvey (Rhodomelaceae, Rhodophyta) in the North Atlantic Ocean highlighted by a DNA barcoding approach. Aquatic Botany 100:67-71.

Global Invasive Species Database.www.issg.org/database. Accessed 2014-10-28.

Gollasch, S. 2006. A new ballast water sampling device for sampling organisms above 50 micron. Aquatic Invasions 1:46-50.

Gollasch, S., M. David, M. Voigt, E. Dragsund, C. L. Hewitt, and Y. Fukuyo. 2007. Critical review of the IMO international convention on the management of ships' ballast water and sediments. Harmful Algae 6:585-600.

Gollasch, S., E. Macdonald, S. Belson, H. Botnen, J. T. Christensen, J. P. Hamer, G. Houvenaghel, A. Jelmert, I. Lucas, D. Masson, T. McCollin, S. Olenin, A. Persson, I. Wallentinus, L. P. M. J. Wetsteyn, and T. Wittling. 2002. Life in Ballast Tanks. Pages 217-231 *in* E. Leppakoski, S. Gollasch, and S. Olenin, editors.Invasive Aquatic Species of Europe.Distribution, Impacts and Management.Springer Netherlands.

Gollasch, S., J. Lenz, M. Dammer, and H.-G. Andres. 2000a. Survival of tropical ballast water organisms during a cruise from the Indian Ocean to the North Sea. Journal of Plankton Research 22:923-937.

Gollasch, S., H. Rosenthal, H. Botnen, J. Hamer, I. Laing, E. Leppakoski, E. Macdonald, D. Minchin, M. Nauke, S. Olenin, S. Utting, M. Voigt, and I. Wallentinus. 2000b. Fluctuations of zooplankton taxa in ballast water during short-term and long-term ocean-going voyages. International Review of Hydrobiology 85:597-608.

Hajibabaei, M., S. Shokralla, X. Zhou, G. A. Singer, and D. J. Baird. 2011. Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. PLoS One 6:e17497.

Harvey, J. B. J., M. S. Hoy, and R. J. Rodriguez. 2009. Molecular detection of native and invasive marine invertebrate larvae present in ballast and open water environmental samples collected in Puget Sound. Journal of Experimental Marine Biology and Ecology 369:93-99

Hebert, P. D. N., A. Cywinska, S. Ball, and J. deWaard. 2003. Biological identifications through DNA barcodes. ProcBiolSci 1512:313-321.

Heger, T. J., E. A. Mitchell, and B. S. Leander. 2013. Holarctic phylogeography of the testate amoeba Hyalospheniapapilio (Amoebozoa: Arcellinida) reveals extensive genetic diversity explained more by

environment than dispersal limitation. Molecular Ecology 22:5172-5184.

Hewitt, C. L., S. Gollasch, and D. Minchin. 2009. The vessel as a vector - biofouling, ballast water and sediments. Pages 117-129 *in* G. Rilov and J. A. Crooks, editors.Biological invasions in marine ecosystems.Springer-Verlag, Berlin Heidelberg.

Hurd, C.L., W.A. Nelson, R. Falshaw and K.F. Neill. 2004. History, current status and future of marine macroalgal research in New Zealand: Taxonomy, ecology, physiology and human uses. Phycological Research 52(2): 80-106.

IMO. 2004. International Convention for the Control and Management of Ships' Ballast Water and Sediments. IMO.

IUCN. 2000. IUCN guidelines for the prevention of biodiversity loss caused by alien invasive species. Approved by the 51st Meeting of the IUCN Council, February 2000. Approved by the 51st Meeting of the IUCN Council, IUCN, Gland, Switzerland.

Ivanova, N. V., T.S. Zemlak, R.H. Hanner and P.D.N. Hebert. 2007. Universal primer cocktails for fish DNA barcoding. Molecular Ecology Notes 7:544–548.

Jerde, C. L., A.R.Mahon, W.L.Chadderton, and D.M. Lodge. 2011. "Sight-unseen" detection of rare aquatic species using environmental DNA. Conservation Letters, 4:150–157.

Jobard, M., S. Rasconi, and T. Sime-Ngando. 2010. Diversity and functions of microscopic fungi: a missing component in pelagic food webs. Aquatic Science 72:255-268.

Jochem, F. J. 1999. Dark survival strategies in marine phytoplankton assessed by cytometric measurement of metabolic activity with fluorescein diacetate. Marine Biology 135:721-728.

Kaczmarska, I., and L. L. Dowe. 1997. Reproductive biology of the red alga *Polysiphonialanosa* (Ceramiales) in the Bay of Fundy, Canada. Marine Biology 128:695-703.

Kelly, R. P., J. A. Port, K. M. Yamahara, R. G. Martone, N. Lowell, P. F. Thomsen, M. E. Mach, M. Bennet, E. Prahler, M. R. Caldwell, and L. B. Crowder. 2014. Harnessing DNA to improve environmental management. Science 344:1455-1456.

King, D. M., and M. N. Tamburri. 2010. Verifying compliance with ballast water discharge regulations. Ocean Development & International Law 41:152-165.

Ko, W.-H. 2003. Long-term storage and survival structure of three species of *Phytophthora* in water. Journal of General Plant Pathology 69(3): 186-188.

Leppakoski, E., S. Gollasch, and S. Olenin. 2002. Invasive aquatic species of Europe - distribution, impact and management. Kluwer Academic Publishers, Dordrecht, Boston, London.

Martin, E., and R. Hines. 2008. Oxford Dictionary of Biology, 6th edition. Oxford University Press.

Maggs, C., and H. Stegenga. 1999. Red algal exotics on North Sea coasts. Helgoland Marine Research 52:243-258.

McCarthy, H. P., and L. B. Crowder. 2000. An overlooked scale of global transport: phytoplankton species richness in ships' ballast water. Biological Invasions 2:321-322.

Meusnier, I., G. A. C. Singer, J-F.Landry, D. A. Hickey, P. D. N. Hebert, and M. Hajibabaei. 2008. A universal DNA mini-barcode for biodiversity analysis. BMC Genomics 9:214.

Minchin, D. 2007. A checklist of alien and cryptogenic aquatic species in Ireland.Aquatic Invasions 2(4): 341-366.

Mountfort, D., K. F. Smith, M. Kirs, J. Kuhajek, J. E. Adamson, and S. A. Wood. 2012. Development of single and multispecies detection methods for the surveillance and monitoring of marine pests in New Zealand. Aquatic Invasions 7:125-128.

Nicholls, K.H.,H.J. MacIsaac. 2004. Euryhaline, sand-dwelling, testate Rhizopods in the Great Lakes. Journal of Great Lakes Research 30: 123-132.

NOBANIS.European Network on Invasive Alien Species.http://www.nobanis.org. Accessed 2014-10-28

Ojaveer, H., B. S. Galil, D. Minchin, S. Olenin, A. Amorim, J. Canning-Clode, P. Chainho, G. H.Copp, S. Gollasch, A.Jelmert, M. Lehtiniemi, C. McKenzie, J. Mikuš, L. Miossec, A.Occhipinti-Ambrogi, M. Pećarević, J. Pederson, G.Quilez-Badia, J.W. Mijsman, A. Zenetos. 2014. Ten recommendations for advancing the assessment and management of non-indigenous species in marine ecosystems. Marine Policy 44:160-165.

Olenin, S., F. Alemany, A. C. Cardoso, S. Gollasch, P. Goulletquer, M. Lehtiniemi, T. McCollin, D. Minchin, L. Miossec, A. Occhipinti-Ambrogi, H. Ojaveer, K. R. Jensen, M. Stankiewicz, I. Wallentinus, and B. Aleksandrov. 2010. Marine Strategy Framework Directive Task Group 2 Report. Non-indigenous species. JRC Scientific and Technical Reports JRC, DG ENV, ICES, 52pp.

Olenin, S., M. Elliott, I. Bysveen, P. F. Culverhouse, D. Daunys, G. B. J. Dubelaar, S. Gollasch, P. Goulletquer, A. Jelmert, Y. Kantor, K. BringsvorMezeth, D. Minchin, A. Occhipinti-Ambrogi, I. Olenina, and J. Vandekerkhove. 2011. Recommendations on methods for the detection and control of biological pollution in marine coastal waters. Marine Pollution Bulletin 62:2598-2604.

Olenin, S., S. Gollasch, S. Jonusas, and I. Rimkute. 2000. En-route investigations of plankton in ballast water on a ship's voyage from the Baltic Sea to the open Atlantic coast of Europe. International Review of Hydrobiology 5-6:577-596.

Olenin, S., D. Minchin, and D. Daunys. 2007. Assessment of biopollution in aquatic ecosystems. Marine Pollution Bulletin 55:379-394.

Olenin, S., A. Narscius, D. Minchin, M. David, B. Galil, S. Gollasch, A. Marchini, A. Occhipinti-Ambrogi, H. Ojaveer, and A. Zaiko. 2014. Making non-indigenous species information systems practical for management and useful for research: an aquatic perspective. Biological Conservation 173:98-107.

Quinn, G. P., and M. J. Keough. 2002. Experimental design and data analysis for biologists. Cambridge University Press, 537 pp.

Pochon, X., N. J. Bott, K. F. Smith, and S. A. Wood. 2013. Evaluating detection limits of Next-Generation Sequencing for the surveillance and monitoring of international marine pests. PLoS One 8:e73935. doi:73910.71371/journal.pone.0073935.

Pochon, X., A. Zaiko, G.A. Hopkins, J. Banks, S.A. Woods. 2015. Characterizing early marine biofouling communities using High-Throughput Sequencing: an assessment of different sampling devices. Biofouling, in press.

Perrigo, A. L., M. Romeralo and S.L. Baldauf. 2012. What's on your boots: an investigation into the role we play in protistdispersal. Journal of Biogeography, 39:998–1003.

Rizzo, D.M., M. Garbelotto and E.A. Hansen. 2005. *Phytophthoraramorum*: Integrative research and management of an emerging pathogen in California and Oregon forests. Annual Review of Phytopathology 43: 309-335.

R-project. 2014. R: A language and environment for statistical computing. in. R Foundation for Statistical Computing, Vienna, Austria, [ISBN 3-900051-07-0], http://www.R-project.org.

Simberloff, D. 2001. Eradication of island invasives: practical actions and results achieved. Trends Ecology Evolution 16:273-274.

Smayda, T. J., and B. Mitchell-Innes. 1974. Dark survival of autotrophic, planktonic marine diatoms. Marine Biology 25:195-202.

Smith, L. D., M. J. Wonham, L. D. McCann, G. M. Ruiz, A. H. Hines, and J. T. Carlton. 1999. Invasion pressure to a ballast-flooded estuary and an assessment of inoculant survival. Biological Invasions 1: 67-87.

Stoof-Leichsenring, K. R., L. S. Epp, M. H. Trauth, and R. Tiedemann. 2012. Hidden diversity in diatoms of Kenyan Lake Naivasha: a genetic approach detects temporal variation. Molecular Ecology 21:1918–1930.

Thomsen, M. S., T. Wernberg, P. Staehr, D. Krause-Jensen, N. Risgaard-Petersen, and B. R. Silliman. 2007. Alien macroalgae in Denmark - a broad-scale national perspective. Marine Biology Research 3:61-72.

Valentini, A., F.Pompanon and P. Taberlet. 2007. DNA barcoding for ecologists. Trends in Ecology and Evolution24:110-117.

Werres, S., R. Marwitz, W.A. Man In't Veld, W.A.M. De Cock, P.J.M. Bonants, M. De Weert, K. Themann, , E. Ilieva, and R.P.Baayen, R.P. 2001. *Phytophthoraramorum*sp. nov., a new pathogen on *Rhododendron* and *Viburnum*. Mycological Research 105:115-1165.

Wilcox TM, McKelvey KS, Young MK, Jane SF, Lowe WH, et al. (2013) Robust Detection of Rare Species Using Environmental DNA: The Importance of Primer Specificity. PLoS ONE 8(3): e59520. doi:10.1371/journal.pone.0059520

Willoughby, L. G. 1965. Some observations on the location of sites of fungal activity at Blelham Tarn.Hydrobiologia 25:352-356.

Wood, S. A., K. F. Smith, J. C. Banks, L. A. Tremblay, L. Rhodes, D. Mountfort, S. C. Cary, and X. Pochon. 2014. Molecular genetic tools for environmental monitoring of New Zealand's aquatic habitats, past, present and the future. New Zealand Journal of Marine and Freshwater Research 47:90-119.

Wonham, M. J., W. C. Walton, G. M. Ruiz, A. M. Frese, and B. S. Galil. 2001. Going to the source: role of the invasion pathway in determining potential invaders. Marine Ecology Progress Series 215:1-12.

Wyatt, T., and J. T. Carlton. 2002. Phytoplankton introductions in European coastal waters: why are so few invasions reported? Pages 41-46 in F. Briand, editor. Alien marine organisms introduced by ships in the Mediterranean and Black Seas. CIESM Workshop Monographs 20, Monaco.

Supp. 1. Results of the taxonomic assignments of the OTUs, resulted from NGS, COI marker. An asterisk designates genera not reported from ballast water surveys earlier (Gollasch et al. 2000a, 2000b; Gollasch et al. 2002; Olenin et al. 2000; Duggan et al. 2005; Flagella et al. 2007; Briski et al. 2012), shaded are genera containing NIS (as reported in AquaNIS, Global Invasive Species Database and NOBANIS).

#OTU	Phyla	Genus	COI1	COI2	COI3
1	Bacillariophyta	*Nitzschia*	110	0	0
2	Bacillariophyta	*Nitzschia*	3	15	0
3	Rhodophyceae	*Dasya**	0	0	15
4	Rhodophyceae	*Polysiphonia**	0	0	10
5	Rhodophyceae	*Plocamium**	0	0	10
6	Oomycota	*Achlya**	0	0	12
7	Oomycota	*Achlya**	0	2	2663
8	Oomycota	*Apodachlya**	0	0	21
9	Oomycota	*Phytophthora**	0	0	14
10	Oomycota	*Phytophthora**	0	0	16
11	Oomycota	*Halophytophthora**	0	0	10
12	Oomycota	*Halophytophthora**	12	2	70
13	Oomycota	*Halophytophthora**	0	0	1729
14	Oomycota	*Pythium**	0	0	11
15	Oomycota	*Pythium**	290	0	3
16	Oomycota	*Pythium**	0	0	14
17	Oomycota	*Pythium**	8	13	13
18	Oomycota	*Pythiogeton**	0	0	35
19	Amoebozoa	*Hyalosphenia**	49	11	845
20	Amoebozoa	*Squamamoeba**	0	0	17
21	Arthropoda	*Unident. copepod*	0	0	17
22	Mollusca	*Dioryx**	0	0	25
23	Mollusca	*Parachondria**	0	0	20
24	Mollusca	*Peringia**	0	1154	0
25	Mollusca	*Pupilla**	0	0	95
26	Mollusca	*Systrophia**	0	0	39
27	Porifera	*Ianthella**	1871	2564	474
28	Rotifera	*Brachionus*	0	0	412
29	Unident. invertebrate		10344	12344	1856

Supp. 2.Results of the taxonomic assignments of the OTUs, resulted from NGS, RuBisCO marker. An asterisk designates algae genera not reported from ballast water surveys earlier (Gollasch et al. 2000b; Gollasch et al. 2002; Olenin et al. 2000; Flagella et al. 2007), shaded are genera containing known NIS (as reported in AquaNIS, Global Invasive Species Database and NOBANIS).

#OTU	Phyla	Genus	RBC1	RBC2	RBC3
1	Bacillariophyta	*Amphora*	12	24	25
2	Bacillariophyta	*Amphora*	6	18	35
3	Bacillariophyta	*Amphora*	3217	2667	2641
4	Bacillariophyta	*Amphora*	38	32	16
5	Bacillariophyta	*Asterionella*	101	54	47
6	Bacillariophyta	*Bacillaria*	1692	958	1234
7	Bacillariophyta	*Berkeleya**	0	1	10
8	Bacillariophyta	*Brockmanniella*	455	187	192
9	Bacillariophyta	*Caloneis*	95	91	113
10	Bacillariophyta	*Campylodiscus**	22	22	29
11	Bacillariophyta	*Campylodiscus**	106	52	45
12	Bacillariophyta	*Cerataulus*	22	0	28
13	Bacillariophyta	*Chaetoceros*	305	0	0
14	Bacillariophyta	*Chaetoceros*	42	30	32
15	Bacillariophyta	*Cyclotella*	36	19	40
16	Bacillariophyta	*Cyclotella*	38	25	22
17	Bacillariophyta	*Cyclotella*	100	92	62
18	Bacillariophyta	*Cyclotella*	7	6	15
19	Bacillariophyta	*Cylindrotheca*	19	29	0
20	Bacillariophyta	*Cylindrotheca*	60	38	66
21	Bacillariophyta	*Cymatopleura**	774	675	750
22	Bacillariophyta	*Climaconeis*	10	1	1
23	Bacillariophyta	*Ctenophora**	562	382	527
24	Bacillariophyta	*Dactyliosolen*	18	0	0
25	Bacillariophyta	*Diatoma*	1720	1113	1504
26	Bacillariophyta	*Endictya**	54	9	8
27	Bacillariophyta	*Entomoneis*	48	40	61
28	Bacillariophyta	*Entomoneis*	68	55	91
29	Bacillariophyta	*Epithemia**	416	245	340
30	Bacillariophyta	*Epithemia**	567	384	471
31	Bacillariophyta	*Eunotia**	13	5	7
32	Bacillariophyta	*Eunotogramma*	0	0	128
33	Bacillariophyta	*Extubocellulus**	64	100	102
34	Bacillariophyta	*Frustulia*	24	30	8
35	Bacillariophyta	*Frustulia*	58	47	66
36	Bacillariophyta	*Gomphonema*	6	6	15
37	Bacillariophyta	*Gomphonema*	90	100	120

38	Bacillariophyta	*Gomphonema*	11	3	9
39	Bacillariophyta	*Gomphonema*	414	329	411
40	Bacillariophyta	*Guinardia*	7	10	12
41	Bacillariophyta	*Haslea**	20	24	28
42	Bacillariophyta	*Hyalodiscus*	64	59	47
43	Bacillariophyta	*Hyalosira**	25	11	13
44	Bacillariophyta	*Hyalosynedra**	307	205	211
45	Bacillariophyta	*Lauderia*	63	33	61
46	Bacillariophyta	*Licmophora*	264	302	418
47	Bacillariophyta	*Lithodesmioides**	113	0	0
48	Bacillariophyta	*Lithodesmium*	13	4	12
49	Bacillariophyta	*Mayamaea**	78	61	95
50	Bacillariophyta	*Mayamaea**	139	92	135
51	Bacillariophyta	*Melosira*	22	0	1
52	Bacillariophyta	*Minidiscus**	0	0	462
53	Bacillariophyta	*Navicula*	170	256	203
54	Bacillariophyta	*Navicula*	2790	2909	3648
55	Bacillariophyta	*Navicula*	1241	1319	1855
56	Bacillariophyta	*Navicula*	2280	2392	2861
57	Bacillariophyta	*Navicula*	24	17	28
58	Bacillariophyta	*Navicula*	18	14	14
59	Bacillariophyta	*Navicula*	0	250	0
60	Bacillariophyta	*Navicula*	5	10	22
61	Bacillariophyta	*Navicula*	6	5	14
62	Bacillariophyta	*Navicula*	251	282	465
63	Bacillariophyta	*Navicula*	1093	1205	940
64	Bacillariophyta	*Navicula*	23	30	26
65	Bacillariophyta	*Navicula*	83	80	164
66	Bacillariophyta	*Navicula*	217	135	111
67	Bacillariophyta	*Navicula*	13	12	8
68	Bacillariophyta	*Navicula*	298	343	364
69	Bacillariophyta	*Neidium**	22	20	20
70	Bacillariophyta	*Nitzschia*	195	166	160
71	Bacillariophyta	*Nitzschia*	28	15	27
72	Bacillariophyta	*Nitzschia*	13	12	13
73	Bacillariophyta	*Nitzschia*	520	272	314
74	Bacillariophyta	*Nitzschia*	28	21	13
75	Bacillariophyta	*Nitzschia*	174	185	141
76	Bacillariophyta	*Nitzschia*	27	48	83
77	Bacillariophyta	*Nitzschia*	0	4	11
78	Bacillariophyta	*Nitzschia*	40	20	46
79	Bacillariophyta	*Nitzschia*	788	721	1157
80	Bacillariophyta	*Odontella*	108	17	138

81	Bacillariophyta	*Odontella*	947	515	1253
82	Bacillariophyta	*Odontella*	78	73	106
83	Bacillariophyta	*Paralia*	68	0	0
84	Bacillariophyta	*Pinnularia*	11	0	32
85	Bacillariophyta	*Pinnularia*	1	3	10
86	Bacillariophyta	*Pinnularia*	5	4	44
87	Bacillariophyta	*Pinnularia*	13	15	21
88	Bacillariophyta	*Pinnularia*	14	12	21
89	Bacillariophyta	*Pinnularia*	50	32	29
90	Bacillariophyta	*Psammodictyon**	18	16	6
91	Bacillariophyta	*Pseudo-nitzschia*	14	12	13
92	Bacillariophyta	*Pseudo-nitzschia*	5	20	11
93	Bacillariophyta	*Pseudosolenia*	90	90	141
94	Bacillariophyta	*Pseudostaurosira**	36	36	0
95	Bacillariophyta	*Pseudostriatella**	1056	646	897
96	Bacillariophyta	*Rhopalodia*	39	37	38
97	Bacillariophyta	*Sellaphora**	38	17	5
98	Bacillariophyta	*Seminavis**	162	351	232
99	Bacillariophyta	*Synedropsis*	5	10	18
100	Bacillariophyta	*Skeletonema*	0	0	50
101	Bacillariophyta	*Skeletonema*	38	65	65
102	Bacillariophyta	*Stauroneis**	146	81	89
103	Bacillariophyta	*Staurosira**	36	34	57
104	Bacillariophyta	*Stephanopyxis*	10	0	0
105	Bacillariophyta	*Striatella*	118	65	79
106	Bacillariophyta	*Surirella*	899	827	1293
107	Bacillariophyta	*Surirella*	22	19	22
108	Bacillariophyta	*Thalassiosira*	1405	118	395
109	Bacillariophyta	*Thalassiosira*	14	13	16
110	Bacillariophyta	*Thalassiosira*	35	51	0
111	Bacillariophyta	*Thalassiosira*	6	5	79
112	Bacillariophyta	*Thalassiosira*	7	7	64
113	Bacillariophyta	*Thalassiosira*	0	1	44
114	Bacillariophyta	*Thalassiosira*	27	107	38
115	Bacillariophyta	*Thalassiosira*	38	58	1
116	Bacillariophyta	*Triceratium*	52	0	58
117	Bacillariophyta	*Trigonium*	20	37	0
118	Bacillariophyta	*Tryblionella**	60	69	41
119	Bacillariophyta	*Tryblionella**	23	9	14
120	Bacillariophyta	*Unident. diatom*	11	14	10
121	Bacillariophyta	*Unident. diatom*	91	94	152
122	Dinophyceae	*Durinskia**	45	19	24
123	Dinophyceae	*Galeidinium**	186	126	148

124	Dinophyceae	*Kryptoperidinium**	818	596	714
125	Dinophyceae	*Peridiniopsis**	197	192	121
126	Dinophyceae	*Peridinium*	36	43	37
127	Phaeophyceae	*Bachelotia**	0	22	10
128	Phaeophyceae	*Chnoospora**	21	24	26
129	Phaeophyceae	*Diplura**	19	12	20
130	Phaeophyceae	*Lobophora**	162	30	31
131	Phaeophyceae	*Stypopodium**	28	22	31
132	Rhodophyceae	*Ceramium**	8	8	33
133	Rhodophyceae	*Delisea**	0	0	30
134	Xanthophyceae	*Botrydiopsis**	2334	2260	2927
135	Xanthophyceae	*Excentrochloris**	52	68	0
136	Xanthophyceae	*Heterococcus**	27	61	1