Marine Environmental Research 112 (2015) 64–72

Contents lists available at ScienceDirect

Marine Environmental Research

journal homepage: www.elsevier.com/locate/marenvrev

Detecting nuisance species using NGST: Methodology shortcomings and possible application in ballast water monitoring

Anastasija Zaiko [a, b, 1], Jose L. Martinez [c, 1], Alba Ardura [a, d], Laura Clusa [d], Yaisel J. Borrell [d], Aurelija Samuiloviene [a, e], Agustín Roca [d], Eva Garcia-Vazquez [d, *]

[a] Marine Science and Technology Center, Klaipeda University, H. Manto 84, 92294 Klaipeda, Lithuania
[b] Coastal and Freshwater Group, Cawthron Institute, 98 Halifax Street East, 7010 Nelson, New Zealand
[c] Sequencing Unit, Edificio Severo Ochoa, C/ Julian Claveria s/n, University of Oviedo, Oviedo 33006, Spain
[d] Department of Functional Biology, C/ Julian Claveria s/n, University of Oviedo, Oviedo 33006, Spain
[e] Department of Ecology and Biology, Klaipeda University, H. Manto 84, 92294 Klaipeda, Lithuania

ARTICLE INFO

Article history:
Received 13 November 2014
Received in revised form
20 June 2015
Accepted 1 July 2015
Available online 6 July 2015

Keywords:
Ballast water
Environmental DNA
Next generation sequencing
Species assignation
Validation

ABSTRACT

Detecting the presence of potential invasive species in ballast water is a priority for preventing their spread into new environments. Next generation sequencing technologies are being increasingly used for exploring and assessing biodiversity from environmental samples. Here we apply high throughput sequencing from DNA extracted from ballast water (BW) samples employing two different platforms, Ion Torrent and 454, and compare the putative species catalogues from the resulting Operational Taxonomic Units (OTU). Water samples were taken from the RV *Polastern* ballast tank in five different days between the second and the twentieth navigation day. Pronounced decrease of oxygen concentration and increase of temperature occurred in the BW during this time, coincident with a progressively higher proportion of unassigned OTU and short reads indicating DNA degradation. Discrepancy between platforms for species catalogues was consistent with previously published bias in AT-rich sequences for Ion Torrent platform. Some putative species detected from the two platforms increased in frequency during the *Polarstern* travel, which suggests they were alive and therefore tolerant to adverse conditions. OTU assigned to the highly invasive red alga *Polysiphonia* have been detected at low but increasing frequency from the two platforms. Although in this moment NGST could not replace current methods of sampling, sorting and individual taxonomic identification of BW biota, it has potential as an exploratory methodology especially for detecting scarce species.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Genetics is increasingly useful for studying marine bioinvasions. Amongst the utilities of DNA, species identification (e.g. Bott et al., 2010; Bucklin et al., 2011), and determination of the origin and time of initial incursion of non-indigenous species (NIS) (e.g. Rius et al., 2014; Teske et al., 2014) are probably the most common. The advent of Next Generation Sequencing Technologies (NGST) and the possibility of analyzing DNA directly from water and sediments (environmental DNA, eDNA) has opened new perspectives for early detection of NIS in marine ecosystems, where invasions might remain unnoticed for extended periods (Freire et al., 2014; Zaiko et al., 2014). Detecting short species-specific eDNA fragments from aquatic environments has theoretically a greater sensitivity over traditional survey methods which can be time-consuming and costly (Ardura et al., 2010; Pochon et al., 2013). Therefore, eDNA analysis is increasingly employed for detection of rare or invasive species (e.g. Ficetola et al., 2008; Rees et al., 2014).

Briefly, NGST coupled with eDNA are based on massive PCR amplification of short DNA sequences from environmental samples (water, sediments, soil, gut contents etc.; Kelly et al., 2014). There are three main different NGST (Shokralla et al., 2012). Illumina has adopted a sequencing-by-synthesis approach, utilizing fluorescently labeled reversible-terminator nucleotides on clonally amplified DNA templates immobilized on the surface of a flow cell. For Ion Torrent and 454 Roche, DNA fragments are clonally amplified by emulsion PCR on the surface of microbeads. Nucleotide

* Corresponding author.
E-mail address: egv@uniovi.es (E. Garcia-Vazquez).
[1] Contributed equally to the article.

http://dx.doi.org/10.1016/j.marenvres.2015.07.002
0141-1136/© 2015 Elsevier Ltd. All rights reserved.

incorporation is detected in Ion Torrent by variation in H^+ concentration, since H^+ is released when a nucleotide is incorporated into a strand of DNA by the polymerase action (PGM). In 454 there is light emission: a nucleotide release generates a pyrophosphate molecule that initiates a series of downstream reactions to produce light by the action of the enzyme luciferase. The sequences obtained are quality filtered and compared with reference databases using generally BLAST methodology. These methodologies produce many thousands of sequences that can be putatively assigned to a species in a relatively short time.

However, these emergent NGST methods are not fully mature for application in environmental surveys yet. Usually there is a delay between the introduction of a new technology and its implementation for practical real-life research. There are very few examples of NGST application in routine monitoring surveys, such as eDNA-based monitoring for detecting the endangered fish vendace in the UK (Winfield et al., 2008). Important drawbacks of these technologies are: high costs (e.g. Bott et al., 2010); lack or ambiguity of reference sequences in databases that hampers species identification, especially for less studied taxa (e.g. Bucklin et al., 2011; Ardura et al., 2013); lack of truly universal primers (Wilcox et al., 2013). There is a trade-off between primers' specificity and the diversity that can be detected, since taxon-specific primers robustly detect target species, but the rest of biodiversity may be overlooked. Another major drawback is that discovering DNA from a species in an environmental sample does not implicate the individuals are alive. Based on mesocosmos experiments it has been proved that DNA molecules can persist in water for a time after the organism is dead and/or has been removed from the experimental setting (Dejean et al., 2011; Thomsen et al., 2012; Piaggio et al., 2013; Pilliod et al., 2014). Partially degraded DNA from cells or remains of metazoans can yield positive PCR amplification of Barcode genes (e.g. Hajibabaei et al., 2006; Valentini et al., 2009). This implies certain pitfalls for NGST application in biosecurity surveys since only living organisms can become invaders and pose risks to the ecosystem. Sequences obtained from dead organisms could be considered as false positives.

Other false positives associated with DNA detection from dead or decaying material are related to nucleotide degradation. The DNA detectability decreases with time after the removal of the species source of DNA (Dejean et al., 2011), and DNA degradation is proportional to biological oxygen demand (or directly to oxygen concentrations), chlorophyll, pH (Barnes et al., 2014) and temperature (Burger et al., 1999; Ficetola et al., 2008). Degraded DNA may cause incorrect species assignment, particularly for phylogenetically related species, when slightly shorter or modified amplicons provide significant hits with related species.

Taken together, false positives and other aforementioned problems could be a serious shortcoming for NGST application in environmental research and monitoring, for instance in ballast water (BW) surveillance. Although the transport of organisms in BW is recognized as one of the main threats to marine biodiversity worldwide (Gollasch et al., 2002; Elliott, 2003), NGST technologies are not applied currently to BW monitoring (Gollasch et al., 2007). Laborious and costly methods based on visual recognition of species are employed instead (King and Tamburri, 2010). Pilot analyses of BW using NGST suggest that the technique, although promising, is still immature because some organisms observed *de visu* are not found in NGST data and vice versa (Zaiko et al., 2015). Experiments are needed on accuracy, robustness and cross-platform validity of NGST results in order to boost application of these novel DNA-based technologies and benefit from their cost-efficiency for BW surveys.

The goal of the present study was to assess the performance of NGST for species detection in BW; the consistency of the eDNA analysis results across NGST platforms; identify possible biases due to DNA decay and/or degradation in BW; and suggest strategies for decreasing the false positive rate. The study was conducted onboard RV Polarstern, as a part of the BW experimental survey.

2. Material and methods

2.1. Ballast water sampling

For the BW experiment, the aft ballast tank (70 m^3) of RV Polarstern was filled with ambient water (13.1 °C and 34 PSU) out of Bremenhaven port on October 28 2012. Five samples for NGS analysis were collected on the 2nd, 4th, 12th, 16th and 21st days of the cruise. For each sample, 100 L of ballast water were pumped through a plankton net (30 cm diameter, 55 μm mesh size). The concentrated material (*ca.* 50 mL) was then vacuum-filtered through 0.12 μm Nuclepore™ membrane, which was thereafter preserved with 96% ethanol until eDNA extraction.

Simultaneously with the sample collection, environmental conditions in the BW were recorded. Temperature, pH and dissolved oxygen concentration were measured with Ysi Professional Plus Multimeter.

2.2. DNA extraction and NGS

The precipitates from membrane filters were removed with sterile blades. DNA was then extracted from the filters using QIAamp DNA Mini Kit (Qiagen) following the manufacturer extraction protocol. The eDNA was quantified by a fluorescence-based method (Picogreen, Invitrogen). The eDNA obtained from each water sample was divided in two aliquots for separate NGS in two different platforms.

A barcode coding region of the cytochrome oxydase sub-unit I gene (COI) was PCR-amplified employing the universal primers miniCOI described by Meusnier et al. (2008). The expected amplicon size of the target region was approximately 145 nucleotides.

The 454 platform was employed for analyzing the five BW samples. PCR amplification and DNA sequencing were carried out by Macrogen (Korea) using a Genome Sequencer FLX (Roche). The GS FLX data processing was performed using the Roche GS FLX software (v2.9). The software uses tag (barcode) sequences to segregate the reads from each sample, by matching the initial and final bases of the reads to the known tag sequences used in the preparation of the libraries.

For cross-platform verification of NGS results, aliquots of Days 2, 12 and 21 samples were sequenced using the platform Ion Personal Genome Machine System (PGM. Lifetechnologies) at the Sequencing Unit of the Oviedo University. Libraries were constructed using the kit Ion Plus Fragment Library Kit (Lifetechnologies) and templates were obtained using the Ion PGM™ Template OT2 200 Kit for AB Library Builder™ System (Lifetechnologies). The templates were loaded on a 314 chip and sequenced using the Ion PGM™ 200 Sequencing Kit v2 (Lifetechnologies). Data processing was performed using the Ion Torrent software (Lifetechnologies)

2.3. Bioinformatics and indicators of DNA degradation

Raw 454 data were processed using PRINSEQ v0.20.4 (Schmieder and Edwards, 2011) for filtering too short or/and too long reads (mode ± 2SD) and also low quality reads (mean ≥ 20). BLAST software was used with e = 0.01 searching for the best hits in the nucleotides database of the NCBI. Assignment of taxonomic classification (best hit) was done using in-house software (Macrogen) obtaining at the end the OTU tables.

Two parameters were considered as indicators of DNA

degradation based on raw reads obtained from 454 NGS: average read length and proportion of unassigned reads. The rationale is that during the degradation process DNA molecules are progressively shorter (read length), and consequently impossible to assign to a known species by BLAST methodology. Therefore, the average read (sequence) length and the proportion of unassigned OTUs are expected to be, respectively, inversely and directly proportional to DNA degradation.

For comparison of taxonomic catalogues retrieved from the two platforms, the two datasets of raw sequences were analyzed employing identical bioinformatics. Sequences were filtered in length (between 130 and 200 bp) and in quality (+20), then BLAST was done (maximum E-value = 0.001, minimum percent identity = 90.0) against NCBI COI sequences using QIIME (Caporaso et al., 2010).

2.4. Expert taxonomic assessment

The presence of DNA from species that cannot survive inside the ballast tank was expected. Initial eDNA diversity of the BW would reflect the biota existing in the ambient waters of the uptake area, including 'external sources' – species inhabiting rivers and estuaries nearby, discharges of the urban and rural sewages, humans, pets, insects, terrestrial and marine birds, aquaculture species, cetaceans and many others. As commented above, it is not necessary that the organisms are present in the water for leaving their DNA traces therein. In order to distinguish true BW organisms from those of external sources, the assigned taxa were revised by experts and only those that have been cited in marine temperate waters are considered as true BW organisms.

All sequences except singletons, that could be taxonomic artifacts from BLAST or mutations, were considered in the expert taxonomic assessment. For conservative approach, we chose genus level assignments for the downstream analysis. The following data were considered for platform comparison: number of assigned BW OTUs (Operational Taxonomic Units, genera level), number of BW OTU counts (number of sequences), proportion of BW OTU counts over total reads (BW OTU counts + external sources + matches with GenBank entries generically identified as "environmental DNA" + matches at higher taxonomic level than genus). In addition, we have counted the number of species and sequence counts matching with recognized invasive species appearing in the ISSG IUCN list and/or described as such in relevant scientific literature.

2.5. Statistical analysis

To verify the differences in taxonomic diversity reported from NGS (454 and Ion Torrent platforms) over the observation period, Principal Component Analysis (PCA) was applied with two-dimensional visualization (ordination of samples and variables against two principal axes). For the PCA, the NGS data (number of sequences per OTU) were pooled to superior taxonomic ranks (Phyla) and standardized by total number per sample.

The composition of taxa shared in two datasets (presence-absence data) was compared across assayed NGS platforms and sampling days using nonmetric multidimensional scaling (NMDS) based on Jaccard similarity matrix. NMDS was undertaken with 100 random restarts and visualized in a two-dimensional plot.

The Wilcoxon signed-rank test for paired samples was applied to compare the summarized results of taxonomic assignments (phylum level) between the two platforms.

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

2.6. Taxon-specific primers for Salmonids and sponges

Taxon-specific primers were designed for PCR amplification of mitochondrial DNA sequences from BW eDNA samples. The target taxa were those providing larger discrepancies between the OTU catalogues obtained from the two platforms: Salmonids and sponges (see Results).

A fragment of the 16S rRNA gene was amplified by PCR, employing the following Salmonidae family specific primers designed in the laboratory:

16S-F-Salm (5′-AAGACCTGTATGAATGGCATC-3′) and 16S-R-Salm (5′-TCGATAGGGACTCTGGGAGA-3′).

They amplify a DNA region of approximately 377 nucleotides located between the sites 2125 and 2502, position sites relative to *Salmo salar* mitochondrion complete genome (GenBank: KF792729.1).

The amplification reaction was performed in a total volume of 20 μl, including Green GoTaq® Buffer 1×, 2 mM MgCl2, 0.25 mM dNTPS, 1 μM of each primer, 2 μl of template DNA and 0.65 U of DNA Taq polymerase (Promega). The PCR conditions were the following: an initial denaturation step at 95 °C for 5 min, 35 cycles at 94 °C for 30 s, annealing at 68 °C for 30 s and elongation at 72 °C for 30 s. And a final step of elongation at 72 °C for 10 min.

For sponges of the class Demospongiae, a fragment of the COI gene was PCR-amplified employing the following family-specific primers designed in the laboratory:

espLO (5′-GGRGCYGGWACKGGKTGRCSGG-3′) and dgHO (5′-TAAACTTCAGGGTGACCAAARAAYCA-3′).

They amplify a DNA region of approximately 370 nucleotides between the positions 349 and 719 of the COI gene sequence of *Ianthella basta* with GenBank accession number JF915543.

The amplification reaction was performed in a total volume of 20 μl, including Green GoTaq® Buffer 1×, 2.5 mM MgCl2, 0.25 mM dNTPS, 1 μM of each primer, 2 μl of template DNA and 0.65 U of DNA Taq polymerase (Promega). The PCR conditions were the following: an initial denaturation step at 95 °C for 5 min, 35 cycles at 95 °C for 1 min, annealing at 48 °C for 1 min and elongation at 72 °C for 1 min. And a final step of elongation at 72 °C for 5 min.

PCR products of the two taxon-specific markers were visualized in 2% agarose gels with 2.5 μl of SimplySafe™. We used DNA extracted from Atlantic salmon (*S. salar*), and from *Hymeniacidon perlevis* and *Haliclona xena* (Demospongiae, orders Suberitida and Haplosclerida respectively) as positive controls for the primers 16S-Salm and esp-LO respectively.

3. Results

3.1. DNA degradation inferred for BW samples

The environmental conditions drastically changed in Polarstern BW between the first and the last sampling days (Fig. 1 top). The main change occurred in oxygen concentration that dropped down to nearly zero. Temperature oscillated, first increasing with a maximum around the Equator (day 14th) then declining further south. A slight pH decrease from 7.8 to 7.2 was also observed.

A total of 122 672 reads were obtained after the relaxed initial quality filter applied to raw NGS 454 data. The quality of these initial NGS reads was measured as: 1) the reduction of average read length over the expected amplicon size (145 nucleotides), and 2) the percent of unassigned OTUs (Fig. 1 bottom). DNA read quality was reduced nearly tenfold for the two parameters. The average amplicon size decreased from 138.3 down to 75.8 nucleotides in the 2nd and 21st days respectively. The percent of unassigned counts increased from 7.05% to 68.07% respectively for the same days. Reduction in molecular data quality was apparently more intense

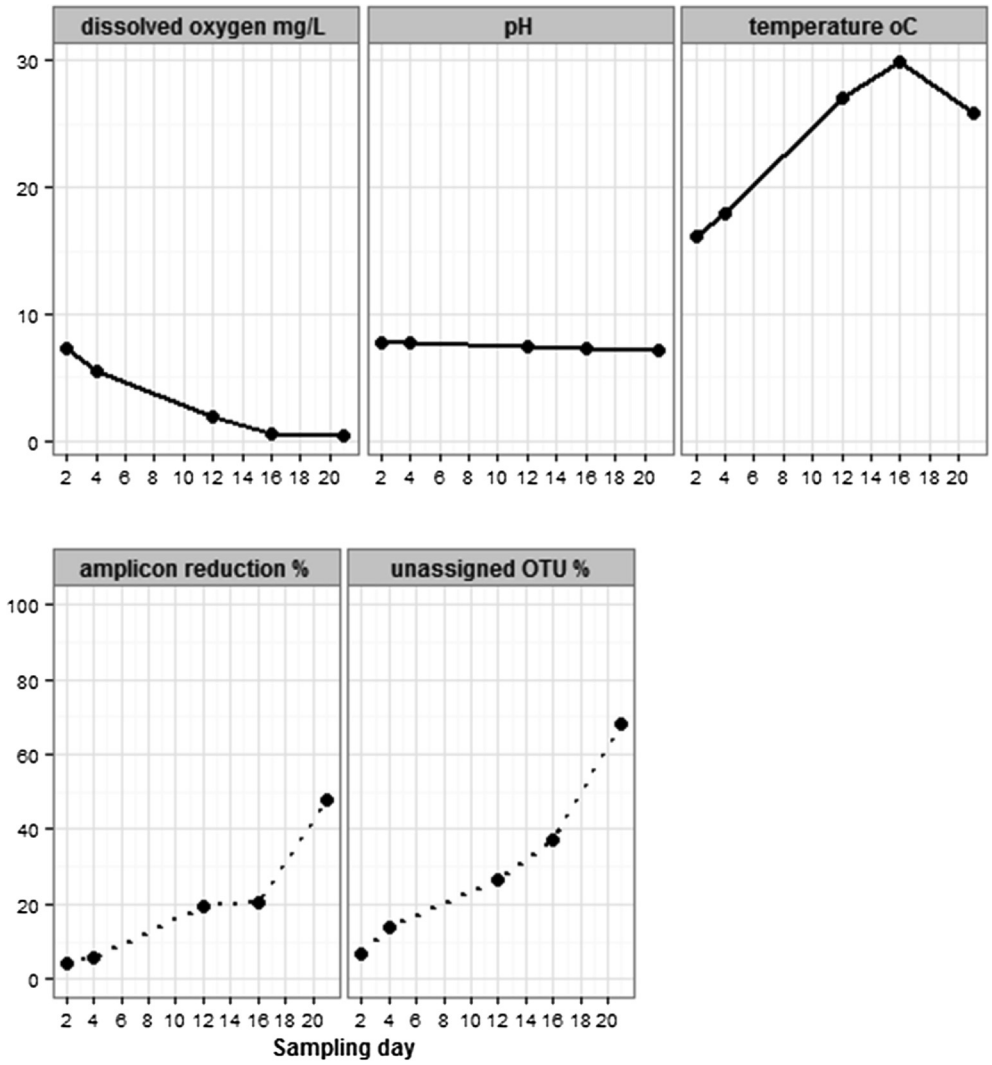

Fig. 1. Physical–chemical parameters of BW in the five sampling days (upper panel) and 454 NGST sequence quality as indicator of DNA degradation (lower panel).

between the 12th and the 21st voyage days, concurrent with the peaks of temperature and anoxia in BW.

3.2. Consistency of taxonomic assignment across NGS platforms

After the bioinformatics treatment with the same stringent filters, the total number of reads yielded from NGS was not very different between the two platforms: 90 955 and 85 913 for the Ion Torrent and 454 platforms respectively (Table 1, Supplementary Table 1). The number of OTU identified from 454 dataset was lower than after the generalized and more relaxed initial quality filter, as expected (122 672 in the relaxed treatment).

External OTUs were assigned to varied genera (data not shown) that cannot live within BW in any developmental stage. These DNA sequences were putatively from humans, dog, cats, amphibians, freshwater algae, insects, arachnidae, marine and terrestrial birds, freshwater and terrestrial molluscs, terrestrial nematodes, land plants and one Cetacean sequences (*Globicephala*, pilot whale). All

Table 1
Summary of OTU identification at genus level for the two NGS platforms assessed (see text for the further details).

	Ion Torrent platform				454 platform					
	Day2	Day12	Day21	Overall	Day2	Day4	Day12	Day16	Day21	Overall
Total BW counts	43768	21923	25264	90955	26950	24175	16116	12730	5942	85913
External sources	16274	7619	17979	41872	815	1834	718	1194	8	4569
Unassigned counts	10505	11272	1705	23482	0	99	0	11	0	110
% Unassigned counts	24.00	51.42	6.75	25.82	0	0.41	0	0.01	0	0.13
BW OTUs (genus level)	42	9	36	71	9	27	2	15	6	31
BW OTUs counts	16989	3032	5580	25601	26135	22242	15398	11525	5934	81234
% BW OTUs counts	33.8	13.83	22.09	28.15	96.97	92	95.54	90.53	99.87	94.55
% dominant OTU counts	15.22	8.64	10.02	26.55	98.65	59.57	95.22	52.22	99.09	39.42
Dominant OTU	*Cleisthenes* (Fish)	*Ianthella* (Porifera)	*Achlya* (Oomycota)	*Cleisthenes* (Fish)	*Salmo* (Fish)	*Salmo* (Fish)	*Salmo* (Fish)	*Salmo* (Fish)	*Salmo* (Fish)	*Salmo* (Fish)

of them could be explained by river and land-based discharges near Bremerhaven, and/or by the presence of some individuals –or biological remains-near the area where the Polarstern tank was filled in.

In addition to those sequences from putative external sources, some sequences could not be assigned down to genus level and were generically grouped as "unassigned counts" (Table 1). The number of unassigned sequences was generally higher in Ion Torrent dataset (on average 47% of all yielded sequences) while in 454 data these did not reach 1% in any of the analyzed samples. The rest of sequences were assigned to genera with species living in marine and/or brackish temperate waters that could potentially survive in BW at least for some hours (true BW OTUs). Those were used for calculating the proportion of sequences corresponding to the dominant species (% dominant OTU counts). In the 454 dataset, the BW sequences obtained on the days 12 and 21 were assigned only to fish species (Supplementary Table 1).

With a similar number of total OTU counts in the two platforms, in the 454 dataset the *Salmo* genus was clearly dominant in the five BW samples analyzed (Table 2), whereas in the Ion Torrent dataset the dominant genus varied from fish on Day 2 to Porifera on Day 12 and Oomycota on Day 21. In this dataset the taxonomic dominance was much less pronounced than in the sequences obtained from 454.

As indicated by the PCA, the 454 samples demonstrated more consistency in terms of biodiversity assigned over the observation period (Fig. 2). However, some taxa (e.g. echinoderms, sponges, rotifers, diatoms) were highly underrepresented in this dataset.

The number of genera identified as true BW (marine, reported from temperate waters) was greater for Ion Torrent than for 454 platform: 71 versus 31 respectively (Table 2). For the Ion Torrent dataset, 39 genera decreased and 31 increased in both % and sequence counts from the first to the 21st day; opposite to this, for the 454 dataset 25 genera decreased and only 3 increased. Genera containing species reported as NIS elsewhere and detected in the 454 dataset were: the copepod *Acartia,* the fish *Ctenogobius*, the fungus *Aphanomyces* and the red algae *Chartransia* and *Polysiphonia* (Supplementary Table 1). In the Ion Torrent dataset the barnacles *Chthamalus* and *Chelonibia*, the fish *Rhinogobius*, the fungus *Saprolegnia* and the red algae *Dasya*, *Polysiphonia* and *Euchema* were found.

When considering the total number of OTU counts at a higher taxonomic level (phylum), the two platforms yielded a clear majority of counts (sequences) of fish and relatively high proportion of oomycetes (Table 2). Main differences between platforms were a much higher percentage of fish (in % counts) and red algae sequences amplified from the 454 platform, while from the Ion Torrent platform Porifera sequences were more abundant. However Wilcoxon signed-rank test showed no statistical significance for the differences of counts yielded per phyla from the two platforms ($P = 0.29$).

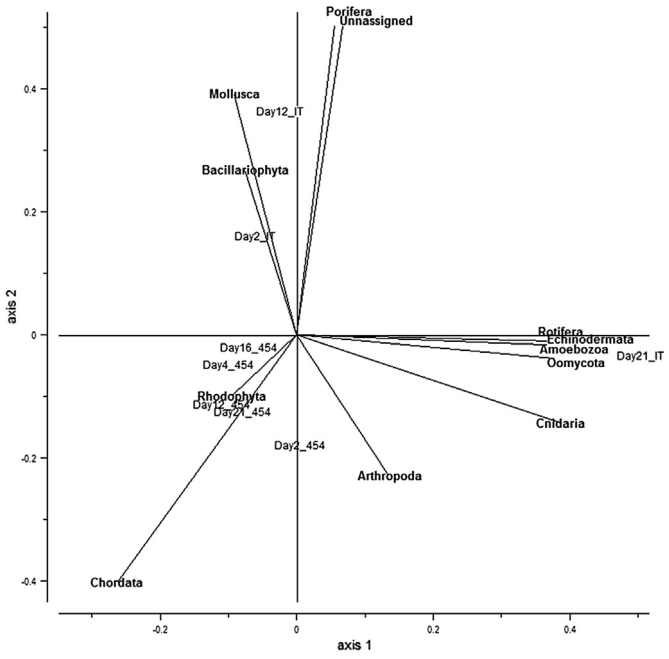

Fig. 2. PCA biplot for NGS data (phylum level, standardized by the total number of sequences per sample) for two NGS platforms. Ion Torrent samples are indicated with label *IT*, 454 data – with label *454*.

The proportions of genera assigned per phylum did not match exactly between platforms, although the difference was not statistically significant (Wilcoxon signed-rank test, $P = 0.96$). In the Ion Torrent platform dataset, Chordata contained more genera than the rest of phyla, while in 454 Rhodophyta was the most diverse phylum (Table 2). The two platforms exhibited, in general, higher diversity within Chordata (fish), Oomycota and Rhodophyta than for the rest of phyla.

3.3. Evolution of putative genera shared from the two platforms

The two platforms employed for NGS analysis provided 20 genera in common (Table 3). One of them, the red alga *Polysiphonia*, contains one of the most aggressive invasive algae, *Polysiphonia brodei*. More fish genera occurred at the beginning of the cruise, while rotifers, red algae and oomycetes genera were more persistent over the observation period. The comparison of shared biodiversity between the two platforms (in terms of presence of assigned genera) showed no discernible pattern, except for a clear match of the Day 2 datasets of the two platforms that were very similar (Fig. 3).

Quantitatively, most of the shared taxa have demonstrated a steady decrease in samples across platforms (Table 3). However some exceptions can be observed. For example, *Tetragonurus* counts fluctuated in the 454 dataset. In the Ion Torrent platform an outbreak of *Peringia* occurred on Day 12, and of *Hyalosphenia* and *Halophytophthora* on Day 21. The number of counts of several oomycetes and the red alga *Polysiphonia* increased in the two platforms over time.

The two most discrepant genera, in number of counts and

Table 2
Summary of the taxonomic assignments based on NGS analysis of the ballast water from the Ion Torrent and 454 platforms, in % over the total number of genera and counts for each phylum. N, total number of genera or OTU counts.

Phylum	Ion Torrent		454	
	% Genera	% Counts	% Genera	% Counts
Annelida	1.41	0.01	0	0
Arthropoda	4.23	0.04	6.06	0.12
Chordata (fish)	43.66	59.28	21.21	84.29
Cnidaria	4.23	0.02	6.06	0.01
Echinoderma	2.82	0.02	0	0
Mollusca	4.23	4.31	6.06	1.36
Porifera	1.41	14.52	3.03	<0.01
Rotifera	2.82	1.51	6.06	0,01
Bacillariophyta	4.23	0.49	0	0
Oomycota	12.68	17.77	15.15	4.72
Rhodophyta	12.68	0.2	30.30	9.34
Amoebozoa	5.63	1.83	3.03	0.15
N	71	25598	33	81234

Table 3
Genera appearing in the datasets from the two NGS platforms assayed, number of yielded sequences respectively.

Kingdom	Phylum	Class	Family	Genus	Ion Torrent			454				
					Day2	Day12	Day21	Day2	Day4	Day12	Day16	Day21
Animalia	Chordata	Actinopterygii	Myctophidae	*Hygophum*	4451	0	44	184	226	0	15	4
			Tetragonuridae	*Tetragonurus*	285	0	0	237	170	52	293	29
			Bothidae	*Bothus*	71	0	0	0	1	0	0	0
			Salmonidae	*Salmo*	295	0	0	24924	14401	15346	6647	5888
			Scorpaenidae	*Scorpaenodes*	121	0	0	0	1	0	0	0
	Cnidaria	Hydrozoa	Hydridae	*Hydra*	1	0	1	9	0	0	0	0
		Scyphozoa	Ulmaridae	*Aurelia*	0	0	1	0	2	0	0	0
	Mollusca	Gastropoda	Hydrobiidae	*Peringia*	0	1100	0	0	748	0	353	0
	Porifera	Demospongiae	Ianthellidae	*Ianthella*	1518	1895	304	1	0	0	0	0
	Rotifera	Monogononta	Brachionidae	*Brachionus*	1	0	384	0	1	0	2	0
				Keratella	0	0	2	0	1	0	0	0
Chromista	Oomycota	Peronospea	Peronosporaceae	*Halophytophthora*	12	2	1546	457	322	0	251	0
			Phytiaceae	*Lagenidium*	1	0	8	2	0	0	0	0
				Pythium	298	15	42	225	1109	0	1008	0
			Saprolegniaceae	*Achlya*	0	2	2532	0	75	0	0	0
Plantae	Rhodophyta	Florideophyceae	Rhodomelaceae	*Polysiphonia*	0	0	16	0	24	0	48	0
			Cystocloniaceae	*Rhodophyllis*	0	0	5	0	1	0	0	0
			Phyllophoraceae	*Schottera*	0	0	1	0	7	0	0	0
			Plocamiaceae	*Plocamium*	0	0	6	0	3	0	0	0
Protozoa	Amoebozoa	Tubulinea	Hyalospheniidae	*Hyalosphenia*	0	0	450	0	78	0	44	0

number of BW samples where they occurred, were the sponge *Ianthella* (Demospongiae) and the fish *Salmo* (Salmonidae) in the Ion Torrent and 454 datasets respectively (Table 3). The sequences obtained for these two OTUs (Supplementary Fig. 1) were clearly different in nucleotide composition (Table 4), with 67.9% and 52.1% AT for *Ianthella* and *Salmo* OTU respectively. In the 5' region the difference was even more accentuated with, respectively, 100% and 73% AT in upstream nucleotides.

PCR amplification using the Salmonidae-specific primers 16S-Sal described above yielded weak but clearly visible bands corresponding to the expected amplicon size from BW DNA of the first, intermediate and last days (Fig. 4A). These results are consistent with Salmo OTUs occurring the three days in the 454 platform. Opposite to this, with the Demospongiae-specific primers only the positive controls give a visible signal of positive PCR amplification (Fig. 4B), suggesting that sponge DNA, if present, was scarce. Again this is coincident with 454 platform results.

4. Discussion

The results found here suggest that NGST can have a value for detection of invasive species from ballast water, but in its current stage of development it is still premature to use Metabarcoding for verification of compliance of BW regulations (King and Tamburri, 2010). Instead it could be considered a complement of the methods of visual taxonomic identification, which on the other hand can be much enhanced if they are coupled with individual Barcoding for taxonomic confirmation (e.g. Valentini et al., 2009). Even a taxonomic quality check of NGS results by an expert seems to be necessary, and automated estimates of biodiversity are not recommended so far if based solely on NGST. A substantial portion of DNA sequences from external sources is expected in BW, particularly during the short voyages or on the first days of observations. Supporting Rees et al. (2014), NGST should not be used to replace or disregard the knowledge and expertise of experienced taxonomists.

As it is apparent from the current results the NGS datasets are not consistent across platforms and a full systematic taxonomic survey using only this novel methodology could not be considered reliable. There are possible biases in Metabarcoding related to peculiarities of the NGS platform and biostatistics pipelines applied, and current trends are to increasingly refine pipelines on a more restricted taxonomic basis (e.g. Balint et al., 2014, for fungi; Liu et al., 2013, for arthropods). Only Day 2 samples from both platforms coincided in terms of taxonomical composition with further increasing discrepancy (higher for Ion Torrent platform). Discrepancy between platforms cannot be explained by different samples analyzed in the present study, since DNA aliquots from the same extraction were taken for analysis in the two different platforms.

A technical issue with the primers cannot be discarded for explaining very high proportion of sequences from dominant taxa in some samples (Table 2). For Deagle et al. (2014) one of the main problems of Metabarcoding based on COI sequences is that COI gene does not contain sufficiently conserved regions for obtaining

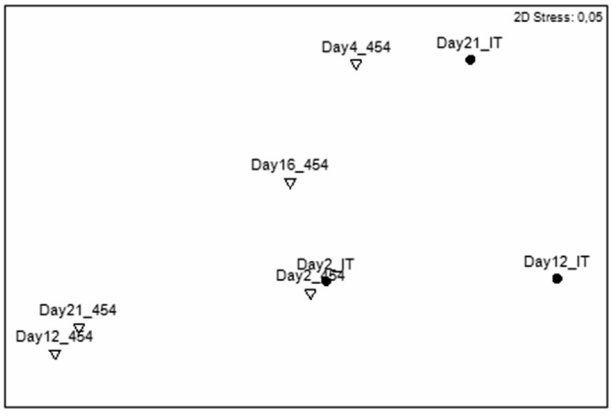

Fig. 3. NMDS plot of the shared genera (from Ion Torrent (*IT*) and 454 (*454*) datasets), based on presence/absence transformed data, Jaccard index similarity matrix.

Table 4
Average nucleotide composition of *Ianthella* and *Salmo* OTUs, in %. Results are given for whole sequences and for the 5' region (15 first nucleotides). N: number of nucleotides.

		N	A	C	G	T
Ianthella	Whole sequence	146	28.8	17.3	14.7	39.1
	5′ region	15	33.3	0	0	66.7
Salmo	Whole sequence	156	23.3	26	21.9	28.8
	5′ region	15	20	20	6.7	53.3

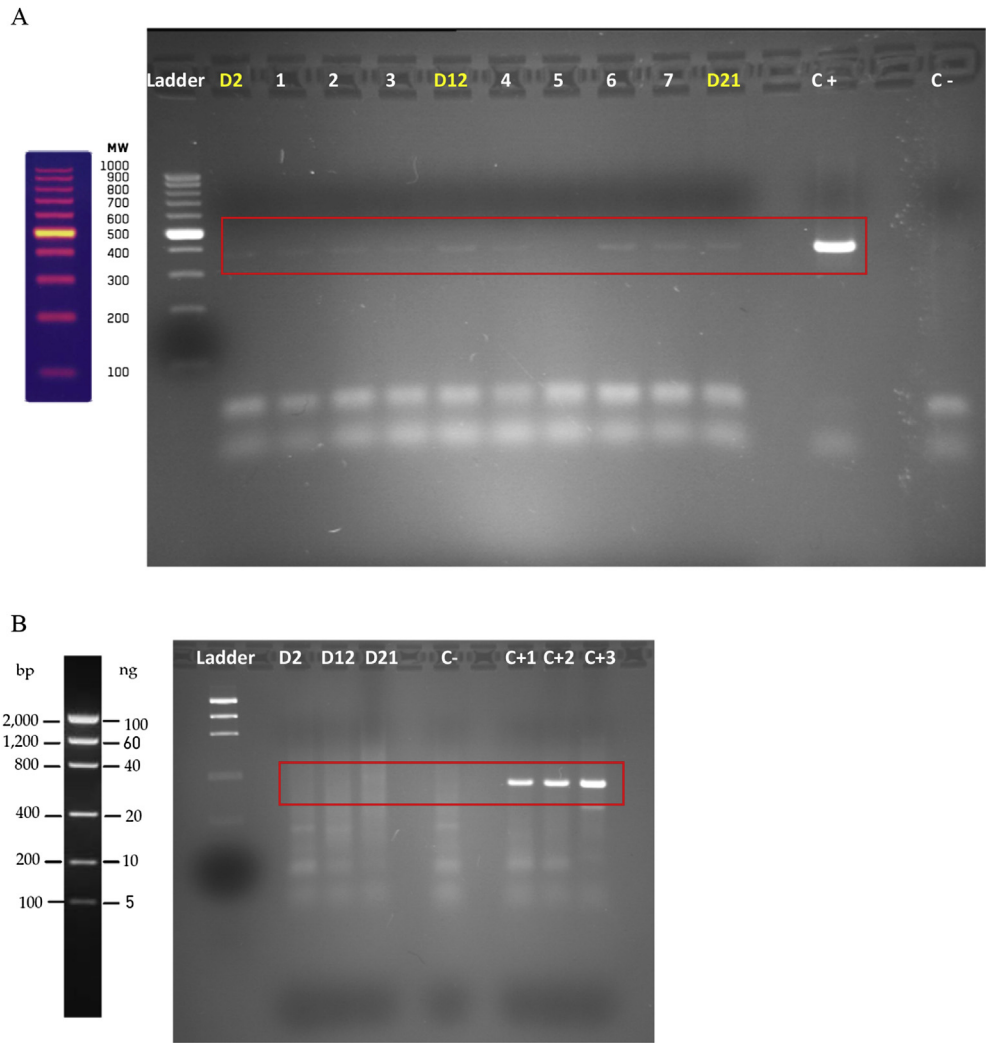

Fig. 4. Agarose gel (2.5%) showing PCR amplification products obtained from ballast water eDNA of different travel days (marked as D2, D12 and D21) with the primers 16S-Salm (A) and espLO (B), specific of Salmonidae and Demospongiae respectively. Positive controls: for Salmonidae primers, C+ = *Salmo salar*; for Demospongiae primers, C+1, C+2 and C+3 are two samples of *H. perlevis* and one of *H. xena* respectively. In (A) PCR products from eDNA extracted from water samples where presence of salmonids was suspected are also included (#1–7); positive amplification with 16S-Sal primers occurred in all cases except #5.

high-quality amplicons from all taxa. Clarke et al. (2014) found that metabarcodes targeting COI introduce taxonomic PCR-amplification bias, amplifying a greater percentage of some species while failing to amplify others. In agreement with Deagle et al. (2014) our results strongly support the need of broadening the catalogue of available markers, and the corresponding reference databases taxonomically verified, for efficient and fully replicable NGST-based BW monitoring and other Metabarcoding applications. Another possibility could be to use PCR-free approaches, such as multiplex sequencing and assembly pipeline (Tang et al., 2014), for avoiding biases due to the different efficiency of primer binding across lineages of organisms.

On the other hand, the bias found here in the composition of the discrepant taxa for AT-rich sequences in the Ion Torrent platform is consistent with data found by other authors in studies of comparison between platforms. In a cross-platform comparison of detection of bacterial pathogens, Frey et al. (2014) found that some sequences were inherently favored in the Ion Torrent platform, whose reads had a higher number of A → T transversions than other platforms. In our study, the dominant *Ianthella* OTUs found from this platform contain a very high proportion of T (double than that of A in the 5' region), suggesting that they were likely overestimated —as also suggested from species-specific primers. High biases in AT-rich sequences from Ion Torrent were also described by Quail et al. (2012) studying the AT-rich genome of *Plasmodium falciparum*. Another artifact described for this platform is early truncation of reads in some specific organisms, attributed also to local sequence characteristics such as homopolymer tracts, local GC content, or local nucleotide composition (Salipante et al., 2014). The results found in our study using taxon-specific primers were more consistent with 454 data than with the Ion Torrent OTU catalogue (Fig. 4), supporting the idea of some biases in the latter benchtop platform. Indeed not finding positive amplification from BW eDNA with the sponge-specific primers does not preclude from sponge DNA being actually present in BW. It may be present in low concentration, as it appears in 454 data (very low proportion of *Ianthella* OTUs detected in the first sample), that would be not consistent with Ion Torrent results. All together, the results suggest that some technical improvements are still necessary, especially in the Ion Torrent platform.

From the OTUs shared in the two platforms, it is clear that BW biota changed significantly towards the end of the travel. This may

reflect higher degradation of DNA from dead individuals that avoids primers from annealing or increase of inhibiting compounds due to the decay processes in the BW. DNA degradation is indicated by increase of unassigned OTUs and shorter NGS reads in 454 dataset, while the total number of yielded sequences decreased. This might imply additional source of error, particularly when employing comparatively short fragment of barcode gene for the analysis (like miniCOI in the current study). Targeting longer DNA fragments (longer amplicons), although might decrease the total number of reads if degraded fragments are excluded for not producing amplicons, could diminish this source of error.

Although we cannot be sure without other type of analysis (see below) the taxa whose proportion increased over time were probably alive; for example outbreaks or sustained increases of a species after several days of navigation could be indicators of living individuals (e.g. Ardura et al., 2015). All the taxa with frequency increments found here have been reported earlier in BW surveys made by conventional methodology (e.g. Carlton and Geller, 1993; Carlton, 1996; Gollasch et al., 2000; Olenin et al., 2000; Briski et al., 2012), therefore NGST can be considered a potential tool to be adjusted in a near future. Of interest is the presence of OTUs assigned to the red alga *Polysiphonia* in the two platforms at increasing frequency over time. This genus contains one of the most aggressive invasive alga species, *Polysiphonia brodiei*, which is included in the ISSG database (http://www.issg.org/database/species/ecology.asp?si=1092&fr=1&sts=tss&lang=EN, accessed November 2014). Native from European and North African waters, it has colonized New Zealand, North America, Japan and Australia, where it is considered one of the 10 most dangerous invaders (Hayes et al., 2005). It has been reported attached on ship hulls and gear (Hewitt et al., 2007; Mineur et al., 2008). Another potentially invasive species of this genus transferred by maritime traffic are *Polysiphonia harveyi* (Hewitt et al., 2007) and *Polysiphonia morrowii*, which is invading South American waters (Raffo et al., 2014). In a review about transfer pathways of seaweeds, Hewitt et al. (2007) reported hull, equipment and gear fouling as preferential pathways for the transfer of *Polysiphonia*. Our results strongly support ballast water as a vector as well. Although increasing in frequency, this genus was not one of the most abundant OTUs found in this study, representing as a maximum 0.28% and 0.42% of counts in Day#21 and #16 samples for Ion Torrent and 454 datasets respectively (Supplementary Table 1). Perhaps just a few cysts were giving the signal. At such a low frequency, they could easily be overlooked in visual counts. Not only because a high taxonomical expertise is needed for identifying small cysts, but also due to the low probability of visually detecting scarce and small cysts under the microscope, and differentiate them from many other more abundant organisms, debris and remains. Their DNA sequences, however, can be detected even when their proportion is small. This could be an example of the utility of NGST for detecting scarce species in BW.

Barcoding from RNA samples (Pawlowski et al., 2014) could serve for solving, at least partially, the problem of differentiating between dead and living individuals. RNA is a fragile molecule of short mean life before degradation, therefore the RNA obtained from a cell —or extracellular-is a recent product of gene expression, a process that happens only in living cells. RNA will thus reflect better the composition of living biodiversity. However this will not solve the problem of discrepant results between different platforms. Contribution from taxonomic experts seems therefore be necessary, although NGST could help as a tool to detect scarce species, and to recommend a more focused examination of otherwise overlooked taxa. In addition, NGST could be improved in different ways for more reliable results that avoid the current discrepancies between platforms. For example, the bias in AT-rich genomes found by Quail et al. (2012) with Ion Torrent was interpreted by the authors as an artifact introduced during amplification; avoidance of library amplification and/or emPCR, or use of more faithful enzymes during emPCR was recommended for eliminating the bias (Quail et al., 2012). Using bidirectional amplicon sequencing and an optimized flow order has also been suggested to minimize sequencing artifacts on this platform (Salipante et al., 2014). Finally, another suggestion is the use of increased throughput chips and library diversity; changes in both key metrics would improve the detection of species present at low levels in a sample (Frey et al., 2014).

5. Conclusions

Discrepancy between platforms for OTU catalogues from aliquots of the same eDNA sample suggests that it is premature to rely solely on NGST for biological BW examination. The proportion of unassigned OTUs as an indicator of DNA degradation could be useful for *en route* BW monitoring. The increase of an OTU with time in BW could be taken as an indicator of being alive, thus be used as an alert of potential invaders. In its present stage, NGST seem to have a value as a complementary tool for BW monitoring, especially for detecting scarce species.

Acknowledgments

We thank the crew of the RV Polarstern for their so valuable collaboration, and an anonymous Reviewer for her/his very helpful comments and ideas for improving the experimental analysis. 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 of the Klaipeda University team was supported by the Research Council of Lithuania Project BALMAN (Development of the ships' ballast water management system to reduce biological invasions. Project # TAP-LLT-14-013).

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.marenvres.2015.07.002.

References

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

Ardura, A., Planes, S., Garcia-Vazquez, E., 2013. Applications of DNA barcoding to fish landings: land authentication and diversity assessments. Zookeys 365, 49–65.

Ardura, A., Zaiko, A., Martinez, J.L., Borrell, Y.J., Garcia-Vazquez, E., 2015. Environmental DNA evidence of transfer of North Sea molluscs across tropical waters. J. Molluscan Stud. http://dx.doi.org/10.1093/mollus/eyv022.

Balint, M., Schmidt, P.A., Sharma, R., Thines, M., Schmitt, I., 2014. An Illumina metabarcoding pipeline for fungi. Ecol. Evol. 4, 2642–2653.

Barnes, M.A., Turner, C.R., Jerde, C.L., Renshaw, M.A., Chadderton, W.L., Lodge, D.M., 2014. Environmental conditions influence eDNA persistence in aquatic systems. Environ. Sci. Technol. 48, 1819–1827.

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

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

Bucklin, A., Steinke, D., Blanco-Bercial, L., 2011. DNA barcoding of marine metazoa. Annu. Rev. Mar. Sci. 3, 471–508.

Burger, J., Hummel, S., Herrmann, B., Henke, W., 1999. DNA preservation: a microsatellite-DNA study on ancient skeletal remains. Electrophoresis 20, 1722–1728.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Gonzalez Pena, A., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T.,

Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., Knight, R., 2010. QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7, 335–336.

Carlton, J.T., 1996. Pattern, process and prediction in marine invasion ecology. Biol. Conserv. 78, 97–106.

Carlton, J.T., Geller, J.B., 1993. Ecological roulette: the global transport of nonindigenous marine organisms. Science 261, 78–82.

Clarke, L.J., Soubrier, J., Weyrich, L.S., Cooper, A., 2014. Environmental metabarcodes for insects: in silico PCR reveals potential for taxonomic bias. Mol. Ecol. Resour. 14 (6), 1160–1170.

Deagle, B.E., Jarman, S.N., Coissac, E., Pompanon, F., Taberlet, P., 2014. DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. Biol. Lett. 10, 20140562.

Dejean, T., Valentini, A., Duparc, A., Pellier-Cuit, S., Pompanon, F., Taberlet, P., Miaud, C., 2011. Persistence of environmental DNA in freshwater ecosystems. PloS One 6, e23398.

Elliott, M., 2003. Biological pollutants and biological pollution – an increasing cause for concern. Mar. Pollut. Bull. 46, 275–280.

Ficetola, G.F., Miaud, C., Pompanon, F., Taberlet, P., 2008. Species detection using environmental DNA from water samples. Biol. Lett. 4 (4), 423–425.

Frey, K.G., Herrera-Galeano, J.E., Redden, C.L., Luu, T.V., Servetas, S.L., Mateczun, A.J., Mokashi, V.P., Bishop-Lilly, K.A., 2014. Comparison of three next-generation sequencing platforms for metagenomic sequencing and identification of pathogens in blood. BMC Genomics 15, 96.

Freire, M., Genzano, G.N., Neumann-Leitao, S., Perez, C.D., 2014. The non-indigenous medusa Blackfordia virginica (Hydrozoa, Leptothecata) in tropical Brazil: 50 years of unnoticed presence. Biol. Invasions 16, 1–5.

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

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

Gollasch, S., Macdonald, E., Belson, S., Botnen, H., Christensen, J.T., Hamer, J.P., Houvenaghel, G., Jelmert, A., Lucas, I., Masson, D., McCollin, T., Olenin, S., Persson, A., Wallentinus, I., Wetsteyn, L.P.M.J., Wittling, T., 2002. Life in ballast tanks. In: Leppakoski, E., Gollasch, S., Olenin, S. (Eds.), Invasive Aquatic Species of Europe. Distribution, Impacts and Management. Springer, Netherlands, pp. 217–231.

Hajibabaei, M., Smith, M.A., Janzen, D.H., Rodriguez, J.J., Whitfield, J.B., Hebert, P.D.N., 2006. A minimalist barcode can identify a specimen whose DNA is degraded. Mol. Ecol. Notes 6, 959–964.

Hayes, K., Sliwa, C., Migus, S., McEnnulty, F., Dunstan, P., 2005. National Priority Pests: Part II, Ranking of Australian Marine Pests. Dept. of the Environment and Heritage CSIRO, Australia, p. 106. http://www.marine.csiro.au/crimp/reports/PrioritypestsFinalreport.pdf (accessed November 2014).

Hewitt, C.L., Campbell, M.L., Schaffelke, B., 2007. Introductions of seaweeds: accidental transfer pathways and mechanisms. Bot. Mar. 50, 326–337.

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

King, D.M., Tamburri, M.N., 2010. Verifying compliance with ballast water discharge regulations. Ocean Dev. Int. Law 41, 152–165.

Liu, S., Li, Y., Lu, J., Su, X., Tang, M., Zhang, R., Zhou, L., Zhou, C., Yang, Q., Ji, Y., Yu, D.W., Zhou, X., 2013. SOAPBarcode: revealing arthropod biodiversity through assembly of Illumina shotgun sequences of PCR amplicons. Methods Ecol. Evol. 4, 1142–1150.

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

Mineur, F., Johnson, M.P., Maggs, C.A., 2008. Macroalgal introductions by Hull fouling on recreational vessels: seaweeds and sailors. Environ. Manag. 42, 667–676.

Olenin, S., Gollasch, S., Jonusas, S., Rimkute, I., 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. Int. Rev. Hydrobiol. 5–6, 577–596.

Pawlowski, J., Esling, P., Lejzerowicz, F., Cedhagen, T., Wilding, T.A., 2014. Environmental monitoring through rotest next-generation sequencing metabarcoding: assessing the impact of fish farming on benthic foraminifera communities. Mol. Ecol. Resour. 14, 1129–1140.

Piaggio, A.J., Engeman, R.M., Hopken, M.W., Humphrey, J.S., Keacher, K.L., Bruce, W.E., Avery, M.L., 2013. Detecting an elusive invasive species: a diagnostic PCR to detect Burmese python in Florida waters and an assessment of persistence of environmental DNA. Mol. Ecol. Resour. 14, 374–380.

Pilliod, D.S., Goldberg, C.S., Arkle, R.S., Waits, L.P., 2014. Factors influencing detection of eDNA from a stream-dwelling amphibian. Mol. Ecol. Resour. 14, 109–116.

Pochon, X., Bott, N.J., Smith, K.F., Wood, S.A., 2013. Evaluating detection limits of next-generation sequencing for the surveillance and monitoring of international marine pests. PloS One 8, e73935, 73910.71371/journal.pone.0073935.

Quail, M.A., Smith, M., Coupland, P., Otto, T.D., Harris, S.R., Connor, T.R., Bertoni, A., Swerdlow, H.P., Gu, Y., 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. BMC Genomics 13, 341.

R-project, 2014. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria [ISBN 3-900051-07-0]. http://www.R-project.org.

Raffo, M.P., Geoffroy, A., Destombe, C., Schwindt, E., 2014. First record of the invasive red alga *Polysiphonia morrowii* Harvey (Rhodomelaceae, Rhodophyta) on the Patagonian shores of the Southwestern Atlantic. Bot. Mar. http://dx.doi.org/10.1515/bot-2013-0056.

Rees, H.C., Maddison, B.C., Middleditch, D.J., Patmore, J.R.M., Gough, K.C., 2014. Review: the detection of aquatic animal species using environmental DNA – a review of eDNA as a survey tool in ecology. J. Appl. Ecol. 51, 1450–1459.

Rius, M., Turon, X., Bernardi, G., Volckaert, F.A.M., Viard, F., 2014. Marine invasion genetics: from spatio-temporal patterns to evolutionary outcomes. Biol. Invasions. http://dx.doi.org/10.1007/s10530-014-0792-0.

Salipante, S.J., Kawashima, T., Rosenthal, C., Hoogestraat, D.R., Cummings, L.A., Sengupta, D.J., Harkins, T.T., Cookson, B.T., Hoffman, N.G., 2014. Performance comparison of Illumina and Ion Torrent Next-Generation Sequencing platforms for 16S rRNA-based bacterial community profiling. Appl. Environ. Microbiol. 80, 7583–7591.

Schmieder, R., Edwards, R., 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics 27 (6), 863–864.

Shokralla, S., Spall, J.L., Gibson, J.F., Hajibabaei, M., 2012. Next-generation sequencing technologies for environmental DNA research. Mol. Ecol. 21, 1794–1805.

Tang, M., Tan, M., Meng, G., Yang, S., Su, X., Liu, S., Song, W., Li, Y., Wu, Q., Zhang, A., Zhou, X., 2014. Multiplex Sequencing of Pooled Mitochondrial Genomes—a Crucial Step toward Biodiversity Analysis Using Mito-metagenomics. http://dx.doi.org/10.1093/nar/gku917. Nucleic Acids Research first published online.

Teske, P.R., Sandoval-Castillo, J., Waters, J.M., Beheregaray, L.B., 2014. Can novel genetic analyses help to identify low-dispersal marine invasive species? Ecol. Evol. 4 (14), 2848–2866.

Thomsen, P.F., Kielgast, J., Iversen, L.L., Moller, P.R., Rasmussen, M., Willerslev, E., 2012. Detection of a diverse marine fish fauna using environmental DNA from seawater samples. PloS ONE 7, e41732.

Valentini, A., Pompanon, F., Taberlet, P., 2009. DNA barcoding for ecologists. Trends Ecol. Evol. 24, 110–117.

Wilcox, T.M., McKelvey, K.S., Young, M.K., Jane, S.F., Lowe, W.H., Whiteley, A.R., Schwartz, M.K., 2013. Robust detection of rare species using environmental DNA: the importance of primer specificity. PloS ONE 8, e59520.

Winfield, I.J., Fletcher, J.M., James, B.J., 2008. Assessment of the Vendace Population of Bassenthwaite Lake Including Observations on Vendace Spawning Grounds. Final Report. Report to Environment Agency, North West Region, and Scottish Natural Heritage. LA/C03462/3, p. 36.

Zaiko, A., Martinez, J.L., Schmidt-Petersen, J., Ribicic, D., Samuiloviene, A., Garcia-Vazquez, E., 2015. Metabarcoding approach for the ballast water surveillance -– an advantageous solution or an awkward challenge? Mar. Pollut. Bull. 92, 25–34.

Zaiko, A., Minchin, D., Olenin, S., 2014. "The day after tomorrow": anatomy of an 'r' strategist aquatic invasion. Aquat. Invasions 9, 145–155.