Under the canopy: community-wide effects of invasive algae in Marine Protected Areas revealed by metabarcoding

Owen S Wangensteen[a], Emma Cebrian[b], Creu Palacín[c], Xavier Turon[d,*]

[a] Ecosystems and Environment Research Centre, School of Environment and Life Sciences, University of Salford, Salford, UK

[b] Aquatic Ecology Institute, University of Girona, Campus Montilivi, Girona, Spain

[c] Department of Evolutionary Biology, Ecology and Environmental Sciences, and Biodiversity Research Institute (IRBio), University of Barcelona, Barcelona, Spain

[d] Center for Advanced Studies of Blanes (CEAB-CSIC), Blanes (Girona), Spain

*Corresponding author.

E-mail address: xturon@ceab.csic.es (X. Turon)

Running title: community-wide effects of invasive algae in MPAs

Highlights

- Hard-bottom communities dominated by invasive algae were studied with metabarcoding
- A size-fractionation method is developed for the metabarcoding of complex samples
- A modified primer set is used for the amplification of a COI fragment
- A semiquantitave abundance index is proposed for statistical analyses
- Results show effects of canopy-forming invasive algae on the underlying communities

Abstract

We analysed with multigene (18S and COI) metabarcoding the effects of the proliferation of invasive seaweeds on rocky littoral communities in two Spanish Marine Protected Areas. The invasive algae studied were *Caulerpa cylindracea*, *Lophocladia lallemandii* and *Asparagopsis armata*. They are canopy-forming, landscape-dominant seaweeds, and we were interested in their effects on the underlying communities of meiobenthos and macrobenthos, separated in two size fractions through sieving. A new semiquantitative treatment of

metabarcoding data is introduced. The results for both markers showed that the presence of the invasive seaweed had a significant effect on the understory communities for *Lophocladia lallemandii* and *Asparagopsis armata* but not for *Caulerpa cylindracea*. Likewise, changes in MOTU richness and diversity with invasion status varied in magnitude and direction depending on the alga considered. Our results showed that metabarcoding allows monitoring of the less conspicuous, but not least important, effects of the presence of dominant invasive seaweeds.

Introduction

Metabarcoding of DNA is emerging as a powerful tool for biodiversity assessment and monitoring (Taberlet et al., 2012; Baird and Hajibabaei, 2012; Bohmann et al., 2014; Thomsen & Willerslev 2015; Aylagas et al., 2016). This technique, albeit still subject to some limitations, will likely become a cornerstone in decision making of management bodies in the near future (Kelly et al., 2014a; Danovaro et al., 2016). In the marine realm, eukaryotic diversity has been analysed using metabarcoding in plankton and sediment communities (reviewed in Carugati et al., 2015; Bucklin et al., 2016; Sinniger et al., 2016). These studies aimed at diverse applications, including community description, beta-diversity patterns, impact assessment, or study of ecological interactions, among others (e.g., Bik et al., 2012; Fonseca et al., 2014; Pawlowski et al., 2016; Brannock et al., 2016; Guardiola et al., 2015, 2016). Less work has been performed on hard-substrate natural communities (e.g., Pearman et al., 2016), which are among the most affected by human activities. The complex nature of these communities, composed of a tri-dimensional array of superimposed strata (from canopy-forming organisms to cryptic microhabitats), poses methodological challenges for the application of metabarcoding techniques (Wangensteen and Turon, 2017).

Metabarcoding has also been used for the study of introduced and invasive marine species. Research in this field followed two different approaches, one focusing on the early detection of particular pest species (targeted or active surveillance; e.g., Ardura et al., 2015; Simpson et al., 2016), and the other involving monitoring of communities for signs of appearance of alien species (passive surveillance; e.g., Comtet et al., 2015; Zaiko et al., 2015; Abad et al., 2016; Brown et al., 2016; Xiong et al., 2016). Another side of invasion biology is the assessment of the impact of alien species on native assemblages. This is usually performed by traditional community analysis methods, involving sampling, sorting, identification, and preparation of qualitative and/or quantitative inventories (e.g., Piazzi et al., 2001; Balata et al., 2004; Box et al., 2010). These tasks are time-consuming, strongly dependent on available taxonomic expertise, and in practice applied only to the analysis of the larger

elements of the fauna and flora which constitute only a minor fraction of the diversity present (Blaxter, 2016). The use of metabarcoding can greatly improve the sensitivity and breadth of the assessment of biodiversity shifts linked to the proliferation of invasive species, but the potential of this approach remains largely unexplored. Metabarcoding allows analysing not only the larger organisms, but also the smaller components of the eukaryotic diversity, likely the first to respond to perturbations and to suffer from cascading events (Schwindt et al., 2001; Gallucci et al., 2012). At the same time, the definition of taxonomic units based on sequence tags allows comparison across spatially and temporally distant studies, which is hardly possible with traditional inventory lists where many taxa are not identified at the species level and are thus in practice unavailable for comparison with other studies.

Among invasive species, seaweeds profoundly alter hard-substrate sublittoral communities, resulting in economic and ecological impacts worldwide (Schaffelke et al., 2006; Williams and Smith, 2007). The effects of the invasive algae are particularly important when they affect benthic habitats harbouring endangered species and long-lived, slow-growing organisms, which are very sensitive to disturbances (Ballesteros, 2006; Casas-Güell et al., 2016). However, although there is increasing concern about the effects of invasive algae on these habitats, they are commonly assessed by measuring changes in the most apparent or emblematic species (often their disappearance). The focus is therefore on measuring lethal effects of the invasion, which are likely irreversible considering the slow dynamics of these habitats (e.g., Cebrian et al., 2012). The integral community-wide study of the habitats undergoing invasion by alien seaweeds afforded by metabarcoding can allow a fine-scale assessment of their effects, both lethal and sub-lethal. At the same time, it provides a tool for monitoring these effects over time, for early detection of alterations, and for follow-up of restoration efforts.

The goal of the present study is to analyse with multigene (18S and COI) metabarcoding the effects of the proliferation of three invasive seaweeds on rocky littoral communities in two Spanish National Parks. Marine reserves have a pivotal role in the conservation of biodiversity, but their performance in the face of non-native species is not well understood (Byers, 2005; Kellner and Hastings, 2009). Evidence to date suggests that reserves are highly vulnerable to invasive species (reviewed in Burfeind et al., 2017) and thus management plans for reserves should include measures to prevent or counteract their impact.

The algae chosen have a big impact in terms of landscape changes (i.e., they are canopy-forming, engineer species). However, changes in the dominant algal species likely imply changes in the understory compartment, and little is known about effects on the smaller components of the community. In other words, does the presence of invasive algae affect the communities "under the canopy"? The small organisms in benthic communities are the most diverse, and likely the first to respond to environmental alteration. We want to showcase the potential of metabarcoding for this kind of studies, detect changes in eukaryotic biodiversity (meio- and macro-organisms), and set the baseline for future monitoring efforts on these communities. To our knowledge, this is the

first time that such questions are addressed by metabarcoding DNA in the context of marine invasion biology.

Material and Methods

Algal communities selected

We have studied macroalgal forest communities invaded by three alien seaweeds: *Lophocladia lallemandii* (Montagne) F. Schmitz and *Caulerpa cylindracea* Sonder in the Mediterranean, and *Asparagopsis armata* Harvey in the Atlantic.

L. lallemandii is probably coming from the Red Sea via the Suez Canal (Verlaque, 1994; Streftaris and Zenetos, 2006). It is currently distributed throughout most of the Mediterranean Sea, covering several types of substrates and homogenizing the appearance of benthic seascapes (Patzner, 1998; Ballesteros et al., 2007; Cebrian and Ballesteros, 2010). *C. cylindracea* is an endemic species from south-western Australia. The mode of introduction in the Mediterranean remains speculative; however, maritime traffic (ballast water and ship hull fouling) and the aquarium trade are the most likely vectors for this high-impact alga. *A. armata* is native to western Australia, this species was probably introduced into European waters through oyster aquaculture. Nowadays it is distributed throughout Europe in both the Atlantic and the Mediterranean shores. All species have a high invasive potential and all of them are included in the black list of invasive species from IUCN (Otero et al., 2013). Both *C. cylindracea* and *L. lallemandii* were recorded for the first time in the study area in 2003 and rapidly spread to almost all benthic communities present between 0 and 45 m depth (Cebrian et al., 2011), while *A. armata* invasion goes back to the late 90's in the area studied (Guiry, 2017).

Sampling

Samples were taken by scuba diving at two Spanish National Parks: the Cabrera Archipelago (Balearic Islands, Northwestern Mediterranean) and the Atlantic Islands of Galicia (Galicia, Northeastern Atlantic) (Fig. S1). Samples of *Lophocladia lallemandii* were collected in Cabrera Island (October 2015) at 10-12 m depth in a vertical wall facing SE, located in the "Imperial" islet (39º07'30.32"N 2º57'37.14"E). *Caulerpa cylindracea* samples were collected at 30-32 m depth in the same wall and dates as the *L. lallemandii* samples. For these algae, replicate samples were collected in areas visually dominated by the invasive seaweed, while the control samples were taken in zones (interspersed with the former) visually free from them. *Asparagopsis armata* was sampled in the Cíes Islands (Galicia) in May 2015 in a shallow community (4-6 m depth) facing E in the "Penela dos Viños" islet (42º12'52.59", 8º52'41.34"W). This community was completely dominated by *A. armata*, mostly in the sporophyte phase (also known as *Falckenbergia rufolanosa*). It was impossible to sample clearly uninvaded zones at this same spot, so the control

samples were taken at the same habitat, depth and orientation in a site in the "Illa do Monteagudo" about 1 Km from the first place (42º13'32.93"N 8º53'51.29W), in a community dominated by *Cystoseira nodicaulis* (Whitering) M. Roberts. Hereafter, we will refer to these datasets (comprising both the invaded and the non-invaded samples) as the Ll (*Lophocladia lallemandii*), Cc (*Caulerpa cylindracea*) and Aa (*Asparagopsis armata*) datasets. Representative images of the communities sampled are shown in Fig. S2.

Sampling followed the protocol described in Wangensteen and Turon (2017). In short, triplicate samples for each condition were obtained by scraping with hammer and chisel quadrats of 25*25 cm to bare rock while SCUBA diving. The samples were placed underwater in plastic bags. Water was eliminated by filtering (63 µm mesh sieve) and replaced with absolute ethanol. Three size fractions (A: > 10 mm; B: 1 – 10 mm; C: 63 µm – 1 mm) were obtained from each replicate sample using a column of sieves. Fraction A (megabenthos *sensu* Rex and Etter 2010) was dominated by the canopy-forming algal species and was not used in this study as the objective was to assess changes in the smaller components of the community. The retained fractions B and C (macrobenthos and meiobenthos, respectively, Rex and Etter 2010) were then homogenized with a blender and stored at -20 ºC until DNA extraction. All equipment was thoroughly washed and cleaned with sodium hypochlorite between successive samples.

DNA processing

DNA was extracted from 10 g of each homogenized sample using PowerMax Soil DNA Isolation Kit (www.mobio.com). Two genes were amplified: a 100-110 bp fragment in the v7 region of the 18S rRNA gene, using the 18S_allshorts primers (Guardiola et al., 2015; forward: 5'-TTTGTCTGSTTAATTSCG-3' and reverse: 5'-TCACAGACCTGTTATTGC-3), and a fragment of the COI gene, amplified with a modification of the forward miCOIintF primer (Leray et al., 2013): 5'-GGWACWRGWTGRACWITITAYCCYCC-3' and the reverse jgHCO2198 primer (Geller et al., 2013): 5'-TAIACYTCIGGRTGICCRAARAAYCA-3'. The forward primer incorporated two more wobble bases and two inosine nucleotides in the most degenerate positions, relative to the original miCOIintF. This was done for increased universality after manually checking the original primer against representative sequences of the main eukaryotic groups. These primers amplified a region of ca. 313 bp, roughly the second half of the "Folmer region" at the 5' end of the COI gene. See in silico analysis and primer logos in Guardiola et al. (2015) for 18S and Wangensteen et al. (2017) for COI. In order to assign sequences to the different samples (or technical controls), 8-base sample-specific tags were attached to the primers. The same tag was used in forward and reverse primers to detect intersample chimeric sequences.

PCR conditions for 18S followed Guardiola et al. (2015). Amplification of COI used AmpliTaq Gold DNA polymerase, with 1 µl of each 5 µM forward and reverse primers, 3 µg of bovine serum albumin and 10 ng of purified DNA in a total volume of 20 µl per sample. The PCR profile included a denaturing step of 10 min at 95 ºC, 35 cycles of 94 ºC 1 min, 45 ºC 1 min and 72 ºC 1 min and a final extension of 5 min at 72 ºC. After PCR, quality of amplifications was

assessed by electrophoresis in agarose gel. All PCR products were purified using Minelute PCR purification kit (www.qiagen.com).

Triplicate technical controls were obtained in two ways: a PCR-blank was amplified using the elution buffer of the DNA isolation kit as a sample. A negative control for the separation protocol was obtained by using a sand sample charred in a muffle furnace at 400 ºC for 24 h. This muffled sand was sieved, extracted and amplified with the same procedures as for samples

All amplification products for each gene were pooled and used to construct two Illumina libraries using the Metafast protocol at Fasteris SA (Plan-les-Ouates, Switzerland, www.fasteris.com). This protocol does not require any further PCR step, thus minimising biases. Each library was sequenced independently in an Illumina MiSeq platform using v3 chemistry (2x150 bp paired-ends for 18S and 2x300 bp paired-ends for COI).

Bioinformatic analyses

The sequences generated were processed using the OBITools package (Boyer et al., 2016). The length of the raw reads was trimmed to a median Phred quality score higher than 30, and paired-reads were assembled using illuminapairedend. The sequences were demultiplexed (using ngsfilter) according to the sample tags. Sequences for which the tags at the two extremes didn't match were discarded. A length filter (obigrep) was applied to the assigned reads (75 – 180 bp for 18S and 300 – 320 bp for COI). The reads were then dereplicated using obiuniq to obtain unique sequences and scanned for chimeras with Uchime (both *de novo* and using the reference databases explained below, Edgar et al., 2011). Clustering in MOTUs (Molecular Operative Taxonomic Units) was made using the Bayesian clustering algorithm implemented in CROP (Hao et al., 2011) with the following parameter sets: l=0.3, u=0.5 for 18S (equivalent to 99% initial clustering) and l=1.5, u=2.5 for COI (equivalent to 95% initial clustering). CROP does not use a hard threshold but rather optimizes iteratively the resulting clusters using a Gaussian mixture model until a finalization criterion is met (Hao et al., 2011).

The most abundant sequence for each MOTU was taken as the representative and was assigned taxonomically with ecotag, which uses a reference database and an explicit phylogeny to assign MOTUs to the last common ancestor (based on the NCBI Taxonomy database) of the chosen sequences. These sequences were defined as the one showing the highest similarity with the query sequence (best-hit sequence), and all other sequences in the database as similar or more to the best hit as the query sequence is. For 18S, the reference database was generated with ecoPCR (Ficetola et al., 2010) from release 117 of the EMBL nucleotide database (27,915 reference sequences from all major eukaryotic groups). For COI, we also used ecoPCR against release 117 of the EMBL plus sequences obtained from the Barcode of Life Datasystems with a custom R script (188,974 reference sequences). Both reference databases are publicly available from http://github.com/metabarpark/Reference-databases. MOTUs that didn't have a match in the database of >0.8 (18S) or >0.75 (COI) similarity were discarded.

Final refining of the dataset consisted of minimal abundance filtering (3 reads minimum). Then, for each MOTU the counts per sample were ordered from lowest to highest and those corresponding to a cumulative frequency of less than 0.03 were set to 0 (this step aimed at eliminating possible cross-sample contamination during the library preparation step). Any MOTU remaining in the negative controls and blanks after the previous step was removed. All MOTUs that were assigned by ecotag to prokaryotes were likewise eliminated. Finally, non-marine organisms were removed (these could be contaminations or DNA of continental origin present in the samples). Additionally, the MOTUs corresponding to the invasive algae themselves were discarded, as we were interested in differences due to other components of the communities and keeping these MOTUs would inflate such differences.

The MOTUs were then classified into the major super-groups of eukaryotes according to Guillou et al. (2013), with one exception: Opisthokonta was split into Metazoa and "other Opisthokonta". Hereafter, when we present results for super-groups, we will include separately Metazoa and "other Opisthokonta", even if they are not strictly super-groups as defined in Guillou et al. (2013). MOTUs that could not be identified at least at the super-group level were eliminated. Metazoan MOTUs were further classified into phyla for additional analyses, and those that could not be assigned a phylum were excluded from analyses within Metazoa. The final datasets, including the representative MOTU sequences, their taxonomic assignment, and the number of reads per sample, are included as supplementary material (Tables S1-S2). The original paired sequences, once quality-filtered, demultiplexed into samples, dereplicated into unique sequences, and chimera scanned, are available from the Mendeley data repository (www.data.mendeley.com/datasets/ DOI… to be filled upon acceptance).

Statistical analyses

Both qualitative and semiquantitative analyses were performed. The former were based on presence/absence of each MOTU per sample. As the quantitative value of metabarcoding data is a matter of debate (see Discussion), we applied a semiquantitative index by ordering the MOTUs of each sample by increasing number of reads and ranking MOTUs from 0 to 4, being 0 absence of the corresponding MOTU and 1-4 a rank indicating that the MOTU falls in the following percentiles of the distribution of ordered MOTUs in the sample: rank 1, ≤50%, rank 2, >50≤75%; rank 3, >75≤90%, rank 4, >90%. In this way, we could minimize potential effects of primer biases, gene copy (or mitochondria) number, and other problems associated with a strictly quantitative interpretation of the data, while at the same time not losing all the information contained in the number of reads. In addition, we used this information exclusively for comparative purposes, and all potential biases should be the same across samples.

Rarefaction curves were obtained with package vegan 2.0-7 for R (Oksanen et al., 2016), to analyse the gain in MOTU richness at increasing numbers of reads for each sample (function *rarecurve)*. Reduced-space graphical representation of the data was obtained by non-metric multidimensional scaling (nMDS)

ordinations. These analyses were performed with the *metaMDS* function of the package vegan with 500 random starts. The nMDS ordinations were obtained for qualitative (presence/absence) and semiquantitative (rank) data by using distance matrices calculated with Jaccard and Bray Curtis coefficients, respectively, in distance form. Procrustes tests (function protest) were used to compare the configurations obtained with the two methods.

Permutational analyses of variance were performed with the Windows PERMANOVA module (Anderson et al., 2008) incorporated in the Primer v6 statistical package (Clarke and Gorley, 2006). Similarity indices based on presence/absence data (Jaccard index) and on ranks of abundance (Bray Curtis index) were used. A three-way PERMANOVA was performed separately for each community type. Condition (invaded, non-invaded) and Fraction (B, C) were considered fixed factors, and Replicate was included as a blocking factor nested under Condition. Tests of multivariate dispersions (PERMDISP) were also run for main factors to ascertain whether significant values in PERMANOVA were a result of different heterogeneity (spread) of the groups instead of different multivariate mean location.

Clusters and heatmaps were obtained with the gplots 3.0.1 package for R (Warnes et al., 2016) using complete-linkage hierarchical clustering of the samples with the semiquantitative data, and SIMPER analyses were performed with PRIMER to identify the MOTUs that contribute the most to the differentiation between invaded and non-invaded communities. These analyses were performed only for the COI data, as more detailed taxonomic assignments can be obtained with this marker.

RESULTS

After quality checks, clustering and elimination of MOTUs below the match or read abundance thresholds, contaminations or those corresponding to non-marine organisms, the final datasets consisted of 2,415 MOTUs and 6,123,192 reads for 18S and 21,184 MOTUs and 5,509,887 reads for COI. A sample of fraction B of one non-invaded Ll (*Lophocladia lallemandii*) replicate had only 2,923 reads for COI and was discarded for further analyses of this gene. Mean number of reads per sample was 154,123 for 18S and 157,342 for COI. All control samples had negligible numbers or reads.

Table 1 lists the number of MOTUs obtained for the three datasets with the two genes, as per eukaryotic super-group and metazoans. Metazoa were the most diverse taxon for both markers, followed by Stramenopiles (18S) and Archaeplastida (COI). Rhizaria and Apusozoa were detected only in the 18S dataset, while Excavata were identified only with COI. Within metazoans, the order of the five most diverse phyla was Arthropoda>Annelida>Cnidaria>Porifera>Bryozoa for 18S, and Arthropoda>Cnidaria>Annelida>Porifera>Mollusca for COI. Brachiopoda, Gnathostomulida, Cephaloryncha and Entoprocta were found only in the 18S

dataset and Ctenophora only in the COI results. The Cc (*Caulerpa cylindracea*) dataset had consistently the highest number of total MOTUs and the Aa (*Asparagopsis armata*) dataset the lowest.

The results of the rarefaction curves for both genes are presented in Fig. S3. It can be seen than the number of MOTUs tend to level off at increasing number of reads for 18S, while a plateau is not evident for the highly diverse COI marker.

The nMDS ordinations of the samples of the three communities are presented for both genes based on qualitative (presence-absence) and semiquantitative (rank abundance) data in Figs. 1-3. For both genes it is apparent that samples cluster according to fraction and to condition (invaded vs non-invaded), except that the latter was not evident in the qualitative MDS for COI in the Cc dataset. The configurations are essentially similar for both presence-absence and rank-abundance data, and the procrustes tests showed a high correlation between them (correlation coefficients >0.889 in all cases, all p<0.001).

PERMANOVA analyses of these datasets (Table 2,3 and Table S3,S4) reveal that the condition (invaded, non-invaded) factor was highly significant in all cases in the Ll and Aa datasets, but not in the case of the Cc communities. Fraction was always significant except for the Ll COI dataset using presence-absence data. The interaction between fraction and condition was not significant in any case. The blocking factor, replicate, was not significant in tests involving qualitative data for both genes. In tests that used semiquantitative information, replicate was always significant except in the Ll dataset for COI. PERMDISP tests were not significant for the condition factor, so significant outcomes for this factor were not due to differences in data heterogeneity. For the fraction factor, PERMDISP tests were significant for the 18S dataset but not in the case of the COI dataset (with the exception of the Cc community analysed with presence-absence data). These outcomes suggest that the effect of invasive algae is species-dependent, that both fractions respond in a similar way to the presence of the invasive species (non-significant interaction term), and that the variability between replicates is particularly important when the quantitative component is taken into account, even if in rank form. In the following, we will present results only for the COI dataset, to take advantage of a more accurate taxonomic identification of MOTUs (see Discussion).

Clustering and heatmap representation of the Bray-Curtis distances for the three communities (Figs. 4-6) showed that for the Cc dataset the samples clustered first by fraction, and then by invasion status. In the Ll and the Aa datasets, however, the samples clustered first by presence or not of the invasive alga and, within each, by fraction. Average between-sample Bray-Curtis distance was lower for the Cc (49.79±0.89%, mean±SE) communities, and higher for the Ll (52.19±1.40%) and Aa (59.39±1.24%) datasets.

We also analysed the effect of the invasion on the diversity parameters of the communities, measured as MOTU richness (rarefied to the number of reads of the sample with less reads to make values comparable), and Shannon's diversity index calculated on the number of reads of each MOTU (relative to total number of reads of each sample to make values comparable). Overall, fraction C was the most MOTU-rich and diverse in all cases (Figs. 7-9). The

results were also community dependent. In the Cc community, values of MOTU richness and diversity were in general higher in the invaded samples, although differences were not significant (*t*-tests). The number of MOTUs increased in the Ll dataset in the presence of the invasive alga in fraction C and decreased in fraction B (in both cases the outcome was significant, *t*-tests). The same trend, albeit not significant, was observed for Shannon's diversity. Finally, in the Aa dataset, the trend was towards a higher diversity in the non-invaded samples, but differences were only significant (*t*-tests) for the Shannon index in fraction B.

The differences in MOTU richness between invaded and non-invaded communities were further analysed taking into account the main groups found in the samples. Changes in number of MOTUs per sample were mainly driven by changes in metazoan richness (Figs. S4-S6). Within metazoans, the most abundant Phyla had a trend of decreasing richness in the non-invaded community for the Cc dataset, and the differences were significant for Cnidaria and Bryozoa (*t*-tests, Fig. S4). Likewise, all major groups of metazoans had lower richness in the non-invaded samples for the Ll dataset, with significant differences in Cnidaria, Porifera, and Chordata (*t*-tests, Fig. S5). In the Aa results, there were overall more metazoan MOTUs in the non-invaded samples, and this was due notably to a significantly higher diversity of molluscs (Fig. S6). Chordata was also significantly more diverse in the non-invaded community, while Echinodermata had significantly more MOTUs per sample in the invaded samples.

The SIMPER analyses identified the MOTUs contributing the most to the differentiation between invaded and non-invaded communities in the three datasets. We took the first 200 of these MOTUs, ordered them by the total number of reads in the corresponding dataset (as an indirect estimator of their importance in the community), and selected the 25 most abundant (in reads) as representing the main discriminating taxa. Information about these MOTUs is given in Tables S5-S7 for the different datasets. It can be seen that most of these discriminating MOTUs corresponded to metazoans, followed by rhodophytes. The taxonomic identification of the MOTUs could be done at different levels of precision (correlated with the percent identity of the best-match in the database, Tables S5-S7). Although in a few cases this percent was low, resulting in unhelpful high-rank assignments, in 60 out of 75 cases the best-hit with the databases was >85% similar, and 30 MOTUs returned a species-level match after the ecotag assignment. Among the top discriminating MOTUs that could be assigned at species or genus level with the reference databases were: *Mesophyllum macroblastum* (Rhodophyta)*, Crambe crambe* (Porifera)*, Clytia linearis* (Cnidaria) and *Elysia viridis* (Mollusca) in the Cc dataset; *Peyssonnelia rubra* (Rhodophyta)*, Cornularia cornucopiae* (Cnidaria)*, Oscarella* sp. (Porifera) and *Cystoseira* sp. (Ochrophyta) in the Ll dataset, and *Dynamene magnitorata* (Arthropoda)*, Callochiton septemvalvis* (Mollusca)*, Pilumnus hirtellus* (Arthropoda) and *Morchellium argus* (Chordata) in the Aa dataset.

DISCUSSION

Our study shows that metabarcoding can be used to analyse the less apparent effects of invasive, landscape-dominant seaweeds, i.e., the impacts on the understory communities. In turn, these impacts are species-dependent; PERMANOVA analyses detected a significant effect of the presence of *Lophocladia lallemandii* and *Asparagopsis armata* on the community composition, but we didn't find significant differences in the communities of macrobenthos and meiobenthos developing in patches dominated by *Caulerpa cylindracea*. Our samples were totally interspersed (Cc and Ll) or spatially close (Aa), in communities at the same depth and orientation, so that environmental variables should not have biased our results and the effects (or lack thereof) found were attributable to the presence of the invasive seaweeds.

The separation in size fractions allowed us to incorporate an important component of variability. Sorting by size seems the best option when there are large biomass differences in the taxa contained in the samples (Wangensteen and Turon, 2017; Elbrecht et al., 2017). The smallest fraction (corresponding to meio-organisms) was the most diverse in terms of MOTU richness and Shannon diversity. Both fractions considered, however, showed the same multivariate trends in response to the presence of invasive algae (non-significant interaction term in PERMANOVA). As for the diversity variables, the trends were also similar across fractions, except in the case of the Ll dataset, were richness and Shannon diversity decreased in fraction B but increased in fraction C in invaded communities.

Although both genes used in the present work told essentially the same story, there were important differences, the most obvious being a large difference in the number of MOTUs detected. Each marker comes with advantages and drawbacks (Bucklin et al., 2016). 18S in general is known to underestimate the diversity present at the species level in metabarcoding studies (in metazoans in particular, Tang et al., 2012; Leray and Knowlton, 2016), and several species or even several genera can share the same sequence for the v7 region amplified (Guardiola et al., 2016). In addition, several fragments of 18S have been commonly used in metabarcoding, each with its own characteristics (Hadziavdic et al., 2014; Tanabe et al., 2016). The fragment of COI sequenced here, on the other hand, is longer and more variable, thus allowing the delimitation of a much higher number of MOTUs. It is likely that this number is higher than the nominal species level richness, given pervasive cryptic diversity in marine organisms (Knowlton, 2000; López-Legentil and Turon, 2005).

18S has been the molecule of choice in metabarcoding studies of eukaryotes so far, but attempts have been made recently to incorporate COI data in these studies (e.g, Leray and Knowlton, 2015; Berry et al., 2015; Aylagas et al., 2016; Elbrecht and Leese, 2017). Being COI the standard molecule in the Barcode of Life initiative for many groups, it seems highly desirable to be able to use the wealth of information that is being presently generated. The main problem of COI is the purported lack of universal primers (Deagle et al., 2014), which can be solved, at least partially, using highly degenerate primers as in the present study. Moreover, assignment at low taxonomic levels (genus, species) is much more reliable with COI than with 18S. A percent match of 97% or higher implies in general a good species match. As a further indication of the taxonomic resolution of the COI marker, 8.21% of MOTUs were shared by our Atlantic and

Mediterranean samples with COI, while this percent rose to an unrealistic 62.57% for 18S, indicating lumping of species in ribosomal DNA MOTUs, at least using the v7 region.

On the other hand, if there are no close sequences in the COI databases, the MOTUs cannot be assigned reliably, even at high ranks. For 18S, assignment at broad taxonomic categories is generally possible. We found more super-groups and more metazoan phyla with 18S than for COI, reflecting the gaps in the database, particularly for the small groups (the non-shared high-level groups are in all cases small-sized organisms). The metazoan MOTUs that could not be assigned to phylum level or lower with 18S were only 8.88% and 9.68% of the total, respectively, for fractions B and C. For COI these figures were much higher: 34.70% (fraction B) and 35.44% (fraction C). In number of reads, however, the identified MOTUs comprised 94.97% and 94.90% (fractions B and C) of the reads for 18S, and 88.24% and 84.16% (fractions B and C) for COI, thus the unassigned groups represented likely a minor fraction of the biomass present.

Metazoa was the most diverse taxon, with Archaeplastida (particularly rhodophytes) being the second group in number of MOTUs for COI. These groups are particularly well represented in the Barcode of Life database of COI, which may influence this outcome, as unassigned MOTUs likely belong to underrepresented taxa. With the 18S results (v7 region), Metazoa was again the most diverse taxon, followed this time by Stramenopiles. In studies targeting particular groups of protists, this should guide the marker choice. For instance, the v4 and v9 regions of the 18S gene have been extensively used for protist metabarcoding in the marine realm (e.g., Pawlowski et al., 2011; Logares et al., 2014; Malviya et al., 2016), with different advantages and pitfalls (Stoeck et al., 2010; Pawlowski et al., 2011). The nature of the target group, marker variability, and completeness of databases should be carefully considered before undertaking any metabarcoding study.

A multigene approach is desirable to obtain the most information from the datasets (Leray and Knowlton, 2016), as markers with different characteristics can be used in parallel. Indeed, the use of several markers in metabarcoding is rapidly gaining momentum (e.g., Clarke et al., 2017; Harvey et al., 2017; Kelly et al., 2017). It is reassuring that the same general conclusions were reached with both markers used here. For management purposes, however, we believe that the higher taxonomic resolution afforded at genus/species level by COI makes it the most desirable marker, even if further database development is obviously necessary to reduce the rate of unassigned MOTUs. However, many of these unassigned genetic tags will surely have a species name in the near future as the databases grow denser.

Besides marker choice and processing protocols, one of the most contentious issues in eukaryote metabarcoding is the quantitative value of the data (Elbrecht and Leese, 2015; Wares and Pappalardo, 2016; Creer et al., 2016). It has been acknowledged that primer bias and differences in copy number of 18S (or mitochondrial abundance in the case of COI) hinder a direct quantitative relationship between number of reads (usually in the form of relative frequency within samples) and biomass of a given MOTU. Nevertheless, most studies

analysing this relationship report a gross correlation, in the sense that more abundant species also tend to be represented by a higher relative or absolute number of reads (reviewed in Lejzerowicz et al., 2015; Barnes and Turner, 2016; Valentini et al., 2016; Bucklin et al., 2016). It can be added here that abundance estimation using traditional methods is not free of biases, either (Shelton et al., 2016). Several metabarcoding studies opted for considering exclusively qualitative information. We consider, however, that there is useful information in the wide range of reads per sample exhibited by the MOTUs, particularly in the context of comparative studies (thus with all biases factored in). A full use of this information requires the analysis of artificial communities to elaborate correction factors (e.g., Thomas et al., 2016), which is not feasible for complex, natural communities. To circumvent this problem, we suggest here a method based on abundance ranks. The semiquantitative or rank-level value of metabarcoding data has been shown in several contexts (e.g., Kelly et al., 2014b; Sun et al., 2015; Abad et al., 2016; Albaina et al., 2016, Hänfling et al., 2016). Our method is expected to filter out potential biases, at the cost of partially sacrificing the quantitative information present. It allows nevertheless incorporating at least a gross estimate of abundance (in biomass).

The three studied seaweeds are considered invasive in the area studied (Boudouresque & Verlaque 2002), having negative effects on the native communities and becoming dominant species. Studies to date have shown their effect on native macroflora and macrofauna (e.g., Piazzi et al., 2001; Balata et al., 2004; Cebrian et al., 2012; Deudero et al., 2010; Marbà et al., 2014) in terms of compositional changes and/or abundance shifts. However, they also have more cryptic impacts. For instance, they can alter food webs and the general functioning of the communities (e.g., Klein and Verlaque, 2010; Alomar et al., 2016; Cabanellas-Reboredo et al., 2010). The three species have also been shown to affect the fitness of herbivorous fish and sea urchins, including commercial species, through the ingestion of chemical compounds from the algae (Tejada et al., 2013; Felline et al., 2014; Magliozzi et al., 2017; Castanho et al., 2017). Thus, a complex pattern of direct and indirect effects emerges; however, few works have addressed the impacts on the smaller components of the communities like the present study. There is evidence that the three species can affect mobile epifaunal components (Box et al., 2010; Guerra-García et al., 2012; Bedini et al., 2015), and in sediment communities *C. cylindracea* can affect meiofaunal composition and sediment characteristics (Pusceddu et al., 2016; Rizzo et al., 2017). To our knowledge, only Bulleri et al. (2017), using an experimental approach, showed that the transplant of *C. cylindracea* to canopy-removal plots did not alter the structure of the understory assemblages of algae and sessile invertebrates, a result in agreement with our findings, albeit based only on visual censuses of macrobenthic taxa. Our study provides evidence on the impacts of these invasive algae on the smaller size fractions of benthic organisms, including for the first time meiofaunal and protist taxa, on hard-substratum communities.

The changes in diversity related to the presence of invasive seaweeds differed among the communities studied. For the Cc dataset no significant change was detected; for the Ll dataset MOTU richness was higher in the non-invaded community for fraction B, but lower for fraction C; in the Aa dataset the trend was towards an increase in diversity values in the non-invaded community.

These results suggest that no generalization is possible, and that the apparent homogenizing effect of the dominant invasive seaweeds can be counteracted by other factors, such as the generation of more microenvironments, changes in species interactions, and increased protection from predation, among others. It has been reported that the establishment of habitat-forming (i.e., engineering) alien species often results in increases of the diversity of the affected communities (Crooks, 2002; Buschbaum et al., 2006), but our results show that this is not always the case, and that the trends can vary according to the species involved and the fraction considered.

The COI results showed a dominance of cnidarians in the Mediterranean samples (Cc and Ll), which were significantly more diverse in invaded communities, while arthropods and molluscs were clearly dominant in the Atlantic samples (Aa dataset), the latter being more diverse in non-invaded communities. The analyses allowed the identification of the species (notably metazoans and red algae) that contributed the most to the differentiation between invaded and non-invaded communities. Further exploitation of the data including detailed analyses of particular groups is beyond the scope of the present article, but indeed the datasets compiled can be used to answer more specific questions (e.g., group-related) as to the composition and differentiation of these communities and fractions.

In conclusion, an exhaustive community characterization using multigene metabarcoding revealed patterns of change in communities that were dominated by invasive seaweeds in marine reserves. These changes were context-dependent and varied with the invasive alga, the fraction analysed, and the eukaryotic group considered. The small-sized components of the benthic communities are the most diverse and most difficult to study with traditional methods. Yet they are of crucial ecological importance and are likely to be the first to respond to perturbations. The use of metabarcoding is a fast and efficient method to study these components and their changes. The acquired datasets represent baseline information for future monitoring of these communities.

COMPETING INTERESTS

The authors declare that there are no competing interests.

Abad D, Albaina A, Aguirre M, Laza-Martínez A, Uriarte U, Iriarte A, Villate F, Estonba A, 2016. Is metabarcoding suitable for estuarine plankton monitoring? A comparative study with microscopy. Marine Biology 163: 149.

Albaina A, Aguirre M, Abad D, Santos M, Estonba A, 2016. 18S rRNA V9 metabarcoding for diet characterization: a critical evaluation with two sympatric zooplanktivorous fish species. Ecology and evolution 6:1809-1824.

Alomar C, Deudero S, Andaloro F, Castriota L, Consoli P, Falautano M, Sinopoli M, 2016. Marine Environmental Research 120:86-92.

Anderson MJ, Gorley RN, Clarke KR, 2008. PERMANOVA for PRIMER: guide to software and statistical methods, Plymouth, UK.

Ardura A, Zaiko A, Martinez JL, Samuiloviene A, Sen D, Garcia-Vazquez E, 2015. eDNA and specific primers for early detection of invasive species - A case study of the bivalve *Rangia cuneata*, currently spreading in Europe. Marine Environmental Research 112:48-55.

Aylagas E, Borja A, Irigoien X, Rodríguez-Ezpeleta N, 2016. Benchmarking DNA metabarcoding for biodiversity-based monitoring and assessment. Frontiers in Marine Science 3:1-12.

Baird DJ, Hajibabaei M, 2012. Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. Molecular Ecology 21:2039-2044.

Balata D, Piazzi L, Cinelli F, 2004. A comparison among assemblages in areas invaded by *Caulerpa taxifolia* and *C. racemosa* on a subtidal Mediterranean rocky bottom. PSZN: Marine Ecology 25:1-13.

Ballesteros E, 2006. Mediterranean coralligenous assemblages: a synthesis of present knowledge. Oceanogr. Marine biology: an Annual Review 44:123-195.

Ballesteros E, Cebrian E, Alcoverro T, 2007. Mortality of shoots of *Posidonia oceanica* following meadow invasion by the red alga *Lophocladia lallemandii*. Botanica Marina 50:8–13.

Bedini R, Bedini M, Bonechi L, Piazzi L, 2015. Effects of non-native turf-forming Rhodophyta on mobile macro-invertebrate assemblages in the north-western Mediterranean Sea. Marine Biology Research 11:430-437.

Barnes MA, Turner CR, 2016. The ecology of environmental DNA and implications for conservation genetics. Conservation Genetics 17:1-17.

Berry O, Bulman C, Bunce M, Coghlan M, Murray DC, Ward RD, 2015. Comparison of morphological and DNA metabarcoding analyses of diets in exploited marine fishes. Marine Ecology Progress Series 540:167-181.

Bik HM, Halanych KM, Sharma J, Thomas WK, 2012. Dramatic shifts in benthic microbial eukaryote communities following the deepwater horizon oil spill. PLoS One 7:e38550.

Blaxter M, 2016. Imagining Sisyphus happy: DNA barcoding and the unnamed majority. Philosophical Transactions of the Royal Society B 371:20150329.

Bohmann K, Evans A, Gilbert MTP, Carvalho GR, Creer S, Knapp M, Yu DW, de Bruyn M, 2014. Environmental DNA for wildlife biology and biodiversity monitoring. Trends in Ecology and Evolution 29:358-367.

Boudouresque CF, Verlaque M, 2002. Biological pollution in the Mediterranean Sea: invasive versus introduced macrophytes. Marine Pollution Bulletin 44:32-38.

Box A, Martin D, Deudero S, 2010. Changes in seagrass polychaete assemblages after invasion by *Caulerpa racemosa* var. *cylindracea* (Chlorophyta: Caulerpales): community structure, trophic guilds and taxonomic distinctness. Scientia Marina 74:317-329.

Boyer F, Mercier C, Bonin A, Le Bras Y, Taberlet P, Coissac E, 2016. OBITOOLS: a UNIX-inspired software package for DNA metabarcoding. Molecular Ecology Resources 16:176-182.

Brannock PM, Ortmann AC, Moss AG, Halanych KM, 2016. Metabarcoding reveals environmental factors influencing spatio-temporal variation in pelagic micro-eukaryotes. Molecular Ecology 25:3593-604.

Brown EA, Chain FJJ, Zhan A, MacIsaac HJ, Cristescu ME, 2016. Early detection of aquatic invaders using metabarcoding reveals a high number of non-indigenous species in Canadian ports. Diversity and Distributions 22:1045-1059.

Bucklin A, Lindeque PK, Rodriguez-Ezpeleta N, Albaina A, Lehtiniemi M, 2016. Metabarcoding of marine zooplankton: prospects, progress and pitfalls. Journal of Plankton Research 38:393-400.

Bulleri F, Benedetti-Cecchi L, Ceccherelli G, Tamburello L, 2017. A few is enough: a low cover of a non-native seaweed reduces the resilience of Mediterranean macroalgal stands to disturbances of varying extent. Biological Invasions 19:2291-2305.

Burfeind DD, Pitt KA, Connolly RM, Byers JE, 2012. Performance of non-native species within marine reserves. Biological Invasions 15:17-28.

Buschbaum C, Chapman AS, Saier B, 2006. How an introduced seaweed can affect epibiota diversity in different coastal systems. Marine Biology 148:743-754.

Byers JE, 2005. Marine reserves enhance abundance but not competitive impacts of a harvested nonindigenous species. Ecology 86:487-500.

Cabanellas-Reboredo M, Blanco A, Deudero S, Tejadda S, 2010. Effects of the invasive macroalga *Lophocladia lallemandii* on the dieet and trophism of *Pinna nobilis* (Mollusca: Bivalvia) and its guests *Pontonia pinnophylax* and *Nepinnotheres pinnotheres* (Crustacea: Decapoda). Scientia Marina 74:101-110.

Carugati L, Corinaldesi C, Dell'Anno A, Danovaro R, 2015. Metagenetic tools for the census of marine meiofaunal diversity: An overview. Marine Genomics 24:11-20.

Casas-Guell E, Cebrian E, Garrabou J, Ledoux JB, Linares C, Teixidó N, 2016. Structure and biodiversity of coralligenous assemblages dominated by the precious red coral *Corallium rubrum* over broad spatial scales. Scientific Reports 6:36535.

Castanho S, Califano G, Soares F, Costa R, Mata L, Pousao-Ferreira P, Ribeiro L, 2017. The effect of live feeds bathed with the red seaweed *Asparagopsis armata* on the survival, growth and physiology status of *Sparus aurata* larvae. Fish Physiology and Biochemistry 43:1043-1054.

Cebrian E. Ballesteros E, 2010. Invasion of Mediterranean benthic assemblages by red alga *Lophocladia lallemandii* (Montagne) F. Schmitz: depth-related temporal variability in biomass and phenology. Aquatic Botany 92:81-85.

Cebrian E, Ballesteros E, Linares C, Tomas F, 2011. Do native herbivores provide resistance to Mediterranean marine bioinvasions? A seaweed example. Biological Invasions. 13:1397-1408.

Cebrian E, Linares C, Marschal C, Garrabou J, 2012. Exploring the effects of invasive algae on the persistence of gorgonian populations. Biological Invasions 14:2647-2656.

Clarke KR, Gorley RN, 2006. Primer v6: user manual/tutorial, E-PRIMER, Plymouth, UK.

Clarke LJ, Beard JM, Swadling KM, Deagle BE, 2017. Effect of marker choice and thermal cycling protocol on zooplankton DNA metabarcoding studies. Ecology and Evolution 2017:1-11.

Comtet T, Sandionigi A, Viard F, Casiraghi M, 2015. DNA (meta)barcoding of biological invasions: a powerful tool to elucidate invasion processes and help managing aliens. Biological Invasions 17:905-922.

Creer S, Deiner K, Frey S, Porazinska D, Taberlet P, Thomas K, Potter C, Bik H, 2016. The ecologist's field guide to sequence-based identification of biodiversity. Methods in Ecology and Evolution 7:1008-1018.

Crooks JA, 2002. Characterizing ecosystem-level consequences of biological invasions: the role of ecosystem engineers. Oikos 97:153-166.

Danovaro R, Carugati L, Berzano M, Cahill AE, Carvalho S, Chenuil A, Corinaldesi C, Cristina S, David R, Dell'Anno A, Dzhembekova N, Garcés E, Gasol JM, Goela P, Féral JP, Ferrera I, Forster RM, Kurekin AA, Rastelli E, Marinova V, Miller PI, Moncheva S, Newton A, Pearman JK, Pitois SG, Reñé A, Rodriguez-Ezpeleta N, Saggiomo V, Simis SGH, Stefanova K, Wilson C, Lo Martire M, Greco S, Cochrane SKJ, Mangoni O, Borja A, 2016. Implementing and innovating marine monitoring approaches for assessing marine environmental status. Frontiers in Marine Science 3, 213.

Deagle BE, Jarman SN, Coissac E, Pompanon F, Taberlet P, 2014. DNA metabarcoding and the cytochrome *c* oxidase subunit I marker: not a perfect match. Biology Letters 10:20140562.

Deudero S, Blanco A, Box A, Mateu-Vicens G, Cabanellas-Reboredo M, Sureda A, 2010. Interaction between the invasive macroalga *Lophocladia lallemandii* and the bryozoan *Retepora grimaldii* at seagrass meadows: density and phyysiological responses. Biological Invasions 12:41-52.

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R, 2011. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27:2194-2200.

Elbrecht V, Leese F, 2015. Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass-sequence relationships with an innovative metabarcoding protocol. PLoS One 10: e0130324.

Elbrecht V, Leese F, 2017. Validation and development of COI metabarcoding primers for freshwater macroinvertebrate bioassessment. Frontiers in Environmental Science 5:11.

Elbrecht V, Peinert B, Leese F, 2017. Sorting things out - assessing effects of unequal specimen biomass on DNA metabarcoding. Ecology and Evolution 7:6918-6926.

Felline S, Mollo E, Ferramosca A, Zara V, Regoli F, Gorbi S, Terlizzi A, 2014. Can a marine pest reduce the nutritional value of Mediterranean fish flesh? Marine Biology 161:1275-1283.

Ficetola GF, Coissac E, Zundel S, Riaz T, Shehzad W, Bessiere J, Taberlet P, Pompanon F, 2010. An *In silico* approach for the evaluation of DNA barcodes. BMC genomics 11:1-10.

Fonseca VG, Carvalho GR, Nichols B, Quince C, Johnson HF, Neill SP, Lambshead JD, Thomas WK, Power DM, Creer S, 2014. Metagenetic analysis of patterns of distribution and diversity of marine meiobenthic eukaryotes. Global Ecology and Biogeography 23:1293-1302.

Gallucci F, Hutchings P, Gribben P, Fonseca G, 2012. Habitat alteration and community-level effects of an invasive ecosystem engineer: a case study along the coast of NSW, Australia. Marine Ecology Progress Series 449:95-108.

Geller J, Meyer C, Parker M, Hawk A, 2013. Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. Molecular Ecology Resources 13:851-861.

Guardiola M, Uriz MJ, Taberlet P, Coissac E, Wangensteen OS, Turon X, 2015. Deep-sea, deep-sequencing: metabarcoding extracellular DNA from sediments of marine canyons. PLoS One 10, e0139633.

Guardiola M, Wangensteen OS, Taberlet P, Coissac E, Uriz MJ, Turon X, 2016. Spatio-temporal monitoring of deep-sea communities using metabarcoding of sediment DNA and RNA. PeerJ 4, e2807.

Guerra-García JM, Ros M, Izquierdo D, Soler-Hurtado MM, 2012. The invasive *Asparagopsis armata* versus the native *Corallina elongata*: differences in associated peracarid assemblages. Journal of Experimental Marine Biology and Ecology 416-417:121-128.

Guillou L, Bachar D, Audic S, Bass D, Berney C, Bittner L, Boutte C, Burgaud G, de Vargas C, Decelle J, del Campo J, Dolan JR, Dunthorn M, Edvardsen B, Holzmann M, Kooistra WHCF, Lara E, Le Bescot N, Logares R, Mahé F, Massana R, Montresor M, Morard R, Not F, Pawlowski J, Probert I, Sauvadet AL, Siano R, Stoeck T, Vaulot D, Zimmermann P, Christen R, 2013. The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. Nucleic Acids Research 41:D597-D604.

Guiry MD, Guiry GM, 2017. AlgaeBase. World-wide electronic publication, National University of Ireland, Galway. http://www.algaebase.org, accessed on 11 March 2017.

Hadziavdic K, Lekang K, Lanzen A, Jonassen I, Thompson EM, Troedsson C, 2014. Characterization of the 18S gene for designing universal eukaryote specific primers. PLoS One 9, e87624.

Hänfling B, Handley LL, Read DS, Hahn C, Li J, Nichols P, Blackman RC, Oliver A, Winfield IJ, 2016. Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. Molecular Ecology 25:3101-3119.

Hao X, Jiang R, Chen T, 2011. Clustering 16S rRNA for OTU prediction: a method of unsupervised Bayesian clustering. Bioinformatics 27:611-618.

Harvey JBJ, Johnson SB, Fisher JL, Peterson WT, Vrijenhoek RC, 2017. Comparison of morphological and next generation DNA sequencing methods for assessing zooplankton assemblages. Journal of Experimental Marine Biology and Ecology 487:113-126.

Kellner JB, Hastings A, 2009. A reserve paradox: introduced heterogeneity may increase regional invasibility. Conservation Letters 2:115–122.

Kelly RP, Port JA, Yamahara KM, Martone RG, Lowell N, Thomsen PF, Mach ME, Bennett M, Prahler E, Caldwell MR, Crowder LB, 2014a. Harnessing DNA to improve environmental management. Science 344:1455-1456.

Kelly RP, Port JA, Yamahara KM, Crowder LB, 2014b. Using environmental DNA to census marine fishes in a large mesocosm. PLoS One 9, e8675.

Kelly RP, Closek CJ, O'Donnell JL, Kralj JE, Shelton AO, Samhouri JF, 2017. Genetic and manual survey methods yield different and complementary views of an ecosystem. Frontiers in Marine Science 3:283.

Knowlton N, 2000. Molecular genetic analyses of species boundaries in the sea. Hydrobiologia 420:73-90.

Lejzerowicz F, Esling P, Pillet L, Wilding TA, Black KD, Pawlowski J, 2015. High-throughput sequencing and morphology perform equally well for benthic monitoring of marine ecosystems. Scientific Reports 5, 13932.

Leray M, Knowlton N, 2015. DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. Proceedings of the National Academy of Sciences of the USA 112:2076-2081.

Leray M, Knowlton N, 2016. Censusing marine eukaryotic diversity in the twenty-first century. Philosophical Transactions of the Royal Society B 371:20150331.

Leray M, Yang JY, Meyer CP, Mills SC, Agudelo N, Ranwez V, Boehm JT, Machida RJ, 2013. A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. Frontiers in Zoology 10:34.

Logares R, Audic S, Bass D, Bittner L, Boutte C, Christen R, Claverie JM, Decelle J, Dolan JR, Dunthorn M, Edvardsen B, Gobet A, Kooistra WHCF, Mahé F, Not F, Ogata H, Pawlowski J, Pernice MC, Romac S, Shalchian-Tabrizi K, Simon N, Stoeck T, Santini S, Siano R, Wincker P, Zingone A, Richards TA, de Vargas C, Massana R, 2014. Patterns of rare and abundant marine microbial eukaryotes. Current Biology 24:813-821.

López-Legentil S, Turon X, 2005. How do morphotypes and chemotypes relate to genotypes? The colonial ascidian *Cystodytes* (Ascidiacea: Polycitoridae). Zoologica Scripta 3:3-14.

Magliozzi L, Almada F, Robalo J, Mollo E, Polese G, Gonçalves EJ, Felline S, Terlizzi A, D'Aniello B, 2017. Cryptic effects of biological invasions: reduction of the aggressive behaviour of a native fish under the influence of an "invasive" biomolecule. PLoS One 12:e0185620.

Malviya S, Scalco E, Audic S, Vincent F, Veluchamy A, Poulain J, Wincker P, Iudicone D, de Vargas C, Bittner L, Zingone A, Bowler C, 2016. Insights into global diatom distribution and diversity in the world's ocean. Proceedings of the National Academy of Sciences of the USA 113: E1516-E1525.

Marbà N, Arthur R, Alcoverro T, 2014. Getting turfed: the population and habitat impacts of *Lophocladia lallemandii* invasions on endemic *Posidonia oceanica* meadows. Aquatic Botany 116:76-82.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H, 2016. Vegan: community ecology package. R-package version 2.3-4. Available at https://cran.r-project.org/package=vegan.

Otero M, Cebrian E, Francour P, Galil B, Savini D. 2013. Monitoring Marine Invasive Species in Mediterranean Marine Protected Areas (MPAs): A strategy and practical guide for managers. IUCN, Malaga, Spain, 136 pp.

Patzner R, 1998. The invasion of *Lophocladia* (Rhodomelaceae, Lophotalieae) at the northern coast of Ibiza (western Mediterranean Sea). Butlletí Societat d'Història Natural de les Balears 41:75–80.

Pawlowski J, Christen R, Lecrop B, Bachar D, Shahbazkia HR, Amaral-Zettler L, Guillou L, 2011. Eukaryotic richness in the abyss: insights from pyrotag sequencing. PLoS One 6:e18169.

Pawlowski J, Esling P, Lejzerowicz F, Cordier T, Visco JA, Martins CIM, Kvalvik A, Staven K, Cedhagen T, 2016. Benthic monitoring of salmon farms in Norway using foraminiferal metabarcoding. Aquaculture Environment Interactions 8:371-386.

Pearman JK, Anlauf H, Irigoien X, Carvalho S, 2016. Please mind the gap - Visual census and cryptic biodiversity assessment at central red sea coral reefs. Marine Environmental Research 118, 20-30.

Piazzi L, Ceccherelli G, Cinelli F, 2001. Threat to macroalgal diversity: effects of the introduced green alga *Caulerpa racemosa* in the Mediterranean. Marine Ecology Progress Series 210:149-159.

Pusceddu A, Fraschetti S, Scopa M, Rizzo L, Danovaro R, 2017. Meiofauna communities, nematode diversity and C degradation rates in seagrass (*Posidonia oceanica* L.) and unvegetated sediments invaded by the algae *Caulerpa cylindracea* (Sonder). Marine Environmental Research 119:88-99.

Rex MA, Ettter RJ, 2010. Deep-sea biodiversity. Pattern and scale. Harvard University Press, Cambridge, Massachusetts.

Rizzo L, Pusceddu A, Stabili L, Alifano P, Fraschetti S, 2017. Potential effects of an invasive seaweed (Caulerpa cylindracea, Sonder) on sedimentary organis matter and microbial metabolic activities. Scientific Reports 7:12113.

Schaffelke B, Smith JE, Hewitt CL, 2006. Introduced macroalgae—a growing concern. Journal of Applied Phycology 18:529–541.

Schwindt E, Bortolus A, Iribarne OO, 2001. Invasion of a Reef-builder Polychaete: Direct and Indirect Impacts on the Native Benthic Community Structure. Biological Invasions 3:137-149.

Shelton AO, O'Donnell JL, Samhouri JF, Lowell N, Williams GD, Kelly RP, 2016. A framework for inferring biological communities from environmental DNA. Ecological Applications 26:1645-1659.

Simpson TJ, Dias PJ, Snow M, Muñoz J, Berry T, 2016. Real-time PCR detection of *Didemnum perlucidum* (Monniot, 1983) and *Didemnum vexillum* (Kott, 2002) in an applied routine marine biosecurity context. Molecular Ecology Resources 17:443–453.

Sinniger F, Pawlowski J, Harii S, Gooday AJ, Yamamoto H, Chevaldonne P, Cedhagen T, Carvalho G, Creer S, 2016. Worldwide analysis of sedimentary DNA reveals major gaps in taxonomic knowledge of deep-sea benthos. Frontiers in Marine Science 3:92.

Stoeck T, Bass D, Nebel M, Christen R, Jones MDM, Breiner HW, Richards TA, 2010. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. Molecular Ecology 19(Suppl.1):21-31.

Streftaris N, Zenetos A, 2006. Alien marine species in the Mediterranean—the 100 "worst invasives" and their impact. Mediterranean Marine Science 7:87–118.

Sun C, Zhao Y, Li H, Dong Y, MacIsaac HJ, Zhan A, 2015. Unreliable quantitation of species abundance based on high-throughput sequencing data of zooplankton communities. Aquatic Biology 24:9-15.

Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E, 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. Molecular Ecology 21:2045-2050.

Tanabe AS, Nagai A, Hida K, Yasuike M, Fujiwara A, Nakamura Y, Takano Y, Katakura S, 2016. Comparative study of the validity of three regions of the 18S-rRNA gene for massively parallel sequencing-based monitoring of the planktonic eukaryote community. Molecular Ecology Resources 16:402-414.

Tang CQ, Leasi F, Obertegger U, Kieneke A, Barraclough TG, Fontaneto D, 2012. The widely used small subunit 18S rDNA molecule greatly underestimates true diversity in biodiversity surveys of the meiofauna. Proceedings of the National Academy of Sciences of the USA 109:16208-16212.

Tejada S, Deudero S, Box A, Sureda A, 2013. Physiological response of the sea urchin Paracentrotus lividus fed with the seagrass Posidonia oceánica and the alien algae Caulerpa racemosa and Lophocladia lallemandii. Marine Environmental Research 83:48-53.

Thomas AC, Deagle BE, Eveson JP, Harsch CH, Trites AW, 2016. Quantitative DNA metabarcoding: improved estimates of species proportional biomass using correction factors derived from control material. Molecular Ecology Resources 16:714-726.

Thomsen PF, Willerslev E, 2015. Environmental DNA - An emerging tool in conservation for monitoring past and present biodiversity. Biological Conservation 183:4-18.

Valentini A, Taberlet P, Miaud C, Civade R, Herder J, Thomsen PF, Bellemain E, Besnard A, Coissac E, Boyer F, Gaboriaud C, Jean P, Poulet N, Roset N, Copp GH, Geniez P, Pont D, Argillier C, Baudoin JM, Peroux T, Crivelli AJ, Olivier A, Acqueberge M, Le Brun M, Moller PR, Willerslev E, Dejean T, 2016. Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. Molecular Ecology 25:929-942.

Verlaque M, 1994. Inventaire des plantes introduites en Méditerranée: origines et repercussions sur l'environnement et les activités humaines. Oceanologica Acta 17:1–23,

Wangensteen OS, Palacín C, Guardiola M, Turon X, 2017. Metabarcoding littoral hard-bottom communities: unexpected diversity and database gaps revealed by two molecular markers. PeerJ Preprints 5, e3429v1. http://dx.doi.org/10.7287/peerj.preprints.3429v1.

Wangensteen OS, Turon X, 2017. Metabarcoding techniques for assessing biodiversity of marine animal forests. In: Marine animal forests. The ecology of benthic biodiversity hotspots (Rossi S, Bramanti L, Gori A, Orejas Saco del Valle C, eds.). Springer International Publishing. Switzerland. pp. 445-473 (Print ISBN: 978-3-319-21011-7).

Wares JP, Pappalardo P, 2016. Can theory improve the scope of quantitative metazoan metabarcoding? Diversity 8:1.

Warnes GR, Bolker B, Bonebakker L, Gentleman R, Huber W, Liaw A, Lumley T, Maechler M, Magnusson A, Moeller S, Schwartz M, Venables B, 2016. gplots: Various R Programming Tools for Plotting Data. R package version 3.0.1. Available at https://cran.r-project.org/package=gplots.

Williams SL, Smith JE, 2007. A global review of the distribution, taxonomy, and impacts of introduced seaweeds. Annual Review of Ecology, Evolution and Systematics 38:327–359.

Xiong W, Li H, Zhan A, 2016. Early detection of invasive species in marine ecosystems using high.throughput sequencing: technical challenges and possible solutions. Marine Biology 163:139.

Zaiko A, Samuiloviene A, Ardura A, García-Vázquez E, 2015. Metabarcoding approach for nonindigenous species surveillance in marine coastal waters. Marine Pollution Bulletin 100:53-59.

Table 1. Number of MOTUs found for the super-groups considered in the different datasets for the 18S and the COI genes. Data are further subdivided by phylum for the Metazoa. Total columns refer to the three datasets pooled. Unassigned means metazoan MOTUs that could not be reliably assigned to any given phylum by the ecotag procedure.

SUPERGROUPS	MOTUs 18S Cc	MOTUs 18S Ll	MOTUs 18S Aa	TOTAL 18S	MOTUs COI Cc	MOTUs COI Ll	MOTUs COI Aa	TOTAL COI
Metazoa	1071	896	705	1534	11275	7278	4088	15982
Stramenopiles	205	172	175	283	1153	543	444	1512
Archaeplastida	161	143	99	227	2200	1829	771	3301
Alveolata	173	133	122	203	95	57	24	108
Rhizaria	44	30	38	63	-	-	-	-
Other Opisthokonta	46	35	24	56	213	126	48	262
Amebozoa	22	16	12	25	2	2	2	2
Hacrobia	12	11	9	16	7	4	2	8
Excavata	-	-	-	-	7	3	3	9
Apusozoa	5	5	6	8	-	-	-	-
TOTAL	1739	1441	1190	2415	14952	9842	5382	21184
METAZOAN PHYLA								
Arthropoda	219	175	169	311	1483	1174	1184	2866
Annelida	176	146	93	238	1080	728	408	1507
Cnidaria	113	127	67	173	2067	1838	263	2561
Porifera	108	95	47	130	943	437	193	1104
Bryozoa	113	69	45	144	814	461	103	903
Nematoda	81	68	49	103	21	22	11	39
Platyhelminthes	60	50	51	95	8	6	22	30
Mollusca	48	37	63	90	366	355	689	909
Chordata	30	29	32	57	264	92	39	324
Nemertea	15	18	14	28	73	78	19	141
Echinodermata	4	8	12	12	9	11	20	29
Xenacoelomorpha	8	5	5	11	57	54	3	64
Brachiopoda	2	1	1	2	-	-	-	-
Gastrotricha	1	2	1	3	1	0	0	1
Rotifera	0	2	2	2	6	7	2	9
Chaetognatha	2	2	0	2	2	2	0	2
Gnathostomulida	1	1	0	1	-	-	-	-
Cephalorhyncha	1	1	0	1	-	-	-	-
Entoprocta	1	0	0	1	-	-	-	-
Ctenophora	-	-	-	-	1	1	0	1
unassigned	88	60	54	130	4080	2012	1132	5492

Table 2. PERMANOVA tests of the factors Condition (invaded or non-invaded, with Replicate as nested blocking factor) and Fraction (B or C) for the Bray-Curtis similarity on semiquantitative data based on 18S. PERMDISP tests of main factors are also included.

Cc dataset	df	SS	Pseudo-F	P-value	Permdisp
Condition	1	1937.5	2.002	0.102	0.174
Fraction	1	4261.7	12.698	0.001	0.002
Replicate (Condition)	4	3871.2	2.884	0.030	
Condition*Fraction	1	326.8	0.974	0.424	
Residual	4	1342.5			

Ll dataset	df	SS	Pseudo-F	P-value	Permdisp
Condition	1	4680.9	4.960	0.002	0.923
Fraction	1	4342.8	12.532	0.001	0.002
Replicate (Condition)	4	943.82	2.724	0.020	
Condition*Fraction	1	540.83	1.561	0.230	
Residual	4	346.54			

Aa dataset	df	SS	Pseudo-F	P-value	Permdisp
Condition	1	4211.4	3.949	0.005	0.402
Fraction	1	4374.1	10.617	0.001	0.001
Replicate (Condition)	4	4266.4	2.589	0.035	
Condition*Fraction	1	752.74	1.827	0.184	
Residual	4	1648			

Table 3. PERMANOVA tests of the factors Condition (invaded or non-invaded, with Replicate as nested blocking factor) and Fraction (B or C) for the Bray-Curtis similarity on semiquantitative data based on COI. PERMDISP tests of main factors are also included

Cc dataset	df	SS	Pseudo-F	P-value	Permdisp
Condition	1	2091.8	1.602	0.191	0.053
Fraction	1	4317.4	9.307	0.001	0.082
Replicate (Condition)	4	5224.3	2.815	0.022	
Condition*Fraction	1	431.1	0.929	0.439	
Residual	4	1855.6			

Ll dataset	df	SS	Pseudo-F	P-value	Permdisp
Condition	1	4015.4	4.341	0.005	0.099
Fraction	1	3664.3	8.831	0.005	0.860
Replicate (Condition)	4	3891.2	2.344	0.102	
Condition*Fraction	1	525.06	1.265	0.348	
Residual	3	1244.8			

Aa dataset	df	SS	Pseudo-F	P-value	Permdisp
Condition	1	6024.2	4.324	0.008	0.855
Fraction	1	5299.6	8.982	0.001	0.338
Replicate (Condition)	4	5573.0	2.361	0.049	
Condition*Fraction	1	696.81	1.181	0.356	
Residual	4	2360.0			

Figure captions

Figure 1. nMDS ordinations of the samples in the *Caulerpa cylindracea* dataset, coded by condition and fraction. Upper row: configurations obtained with 18S; lower row: configurations obtained with COI. Results based on Jaccard index (J, presence-absence data) and Bray-Curtis index (BC, semiquantitative data). Numbers in the lower right corners indicate stress of the final configurations.

Figure 2. nMDS ordinations of the samples in the *Lophocladia lallemandii* dataset, coded by condition and fraction. Upper row: configurations obtained with 18S; lower row: configurations obtained with COI. Results based on Jaccard index (J, presence-absence data) and Bray-Curtis index (BC, semiquantitative data). Numbers in the lower right corners indicate stress of the final configurations.

Figure 3. nMDS ordinations of the samples in the *Asparagopsis armata* dataset, coded by condition and fraction. Upper row: configurations obtained with 18S; lower row: configurations obtained with COI. Results based on Jaccard index (J, presence-absence data) and Bray-Curtis index (BC, semiquantitative data). Numbers in the lower right corners indicate stress of the final configurations.

Figure 4. Heatmap representation of the clusters and Bray-Curtis distance matrix based on semiquantitative COI data for the Cc dataset. In sample names, Inv/Ninv denote samples with or without the invasive algae, the number identifies the replicate, and the final letter refers to the fraction (B or C). Vertical colored bars code for fraction (blue: fraction B, turquoise: fraction C); horizontal colored bars code for invasion status (green: non-invaded, red: invaded).

Figure 5. Heatmap representation of the clusters and Bray-Curtis distance matrix based on semiquantitative COI data for the Ll dataset. In sample names, Inv/Ninv denote samples with or without the invasive algae, the number identifies the replicate, and the final letter refers to the fraction (B or C). Vertical colored bars code for fraction (blue: fraction B, turquoise: fraction C); horizontal colored bars code for invasion status (green: non-invaded, red: invaded).

Figure 6. Heatmap representation of the clusters and Bray-Curtis distance matrix based on semiquantitative COI data for the Aa dataset. In sample names, Inv/Ninv denote samples with or without the invasive algae, the number identifies the replicate, and the final letter refers to the fraction (B or C). Vertical colored bars code for fraction (blue: fraction B, turquoise: fraction C); horizontal colored bars code for invasion status (green: non-invaded, red: invaded).

Figure 7. COI MOTU richness (rarefied to the size of the sample with the least number of reads) and Shannon's diversity index for each condition and fraction of the Cc dataset.

Figure 8. COI MOTU richness (rarefied to the size of the sample with the least number of reads) and Shannon's diversity index for each condition and fraction of the LI dataset. Asterisks indicate significant differences between invaded and non-invaded communities.

Figure 9. COI MOTU richness (rarefied to the size of the sample with the least number of reads) and Shannon's diversity index for each condition and fraction of the Aa dataset. Asterisks indicate significant differences between invaded and non-invaded communities.
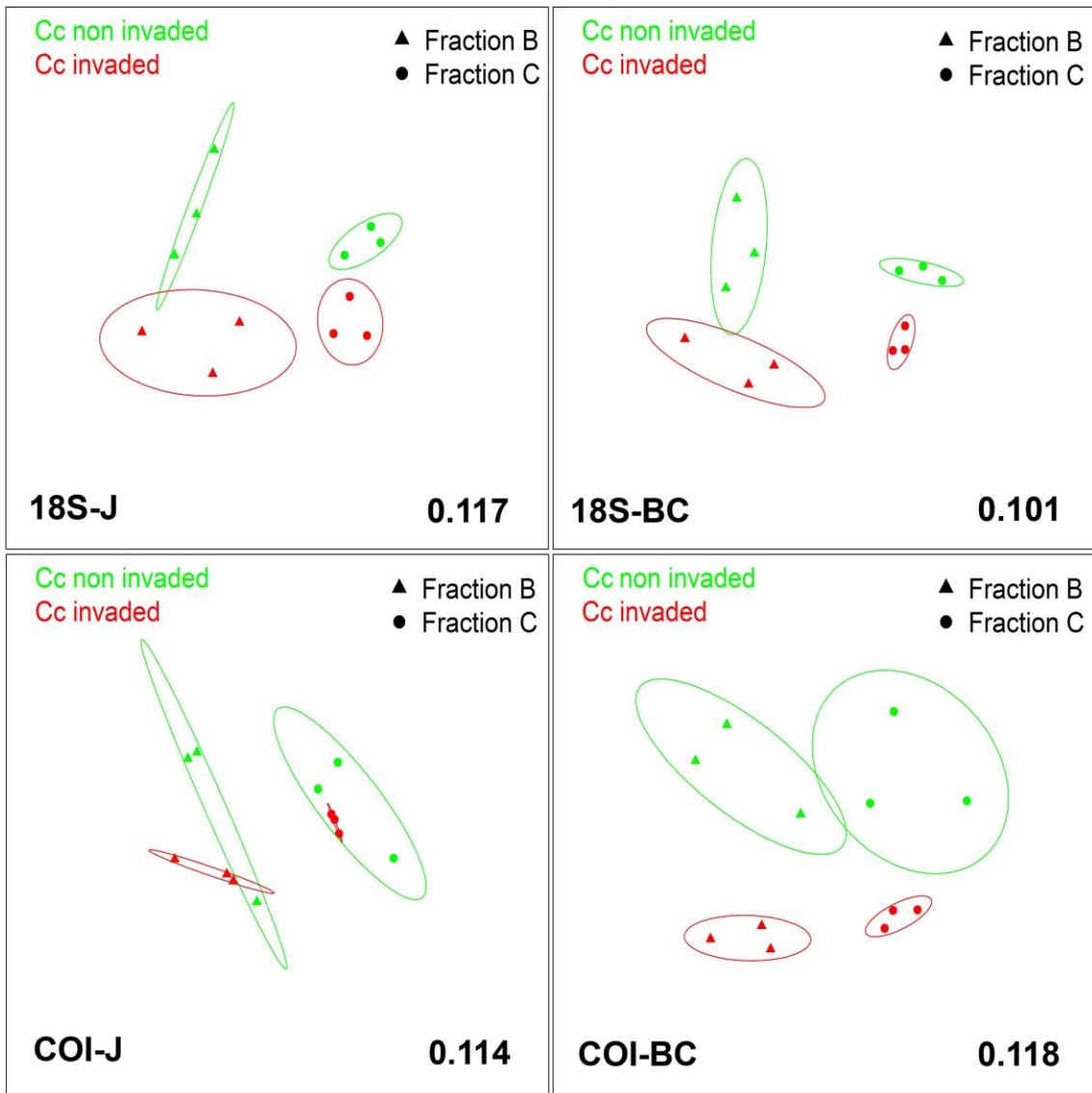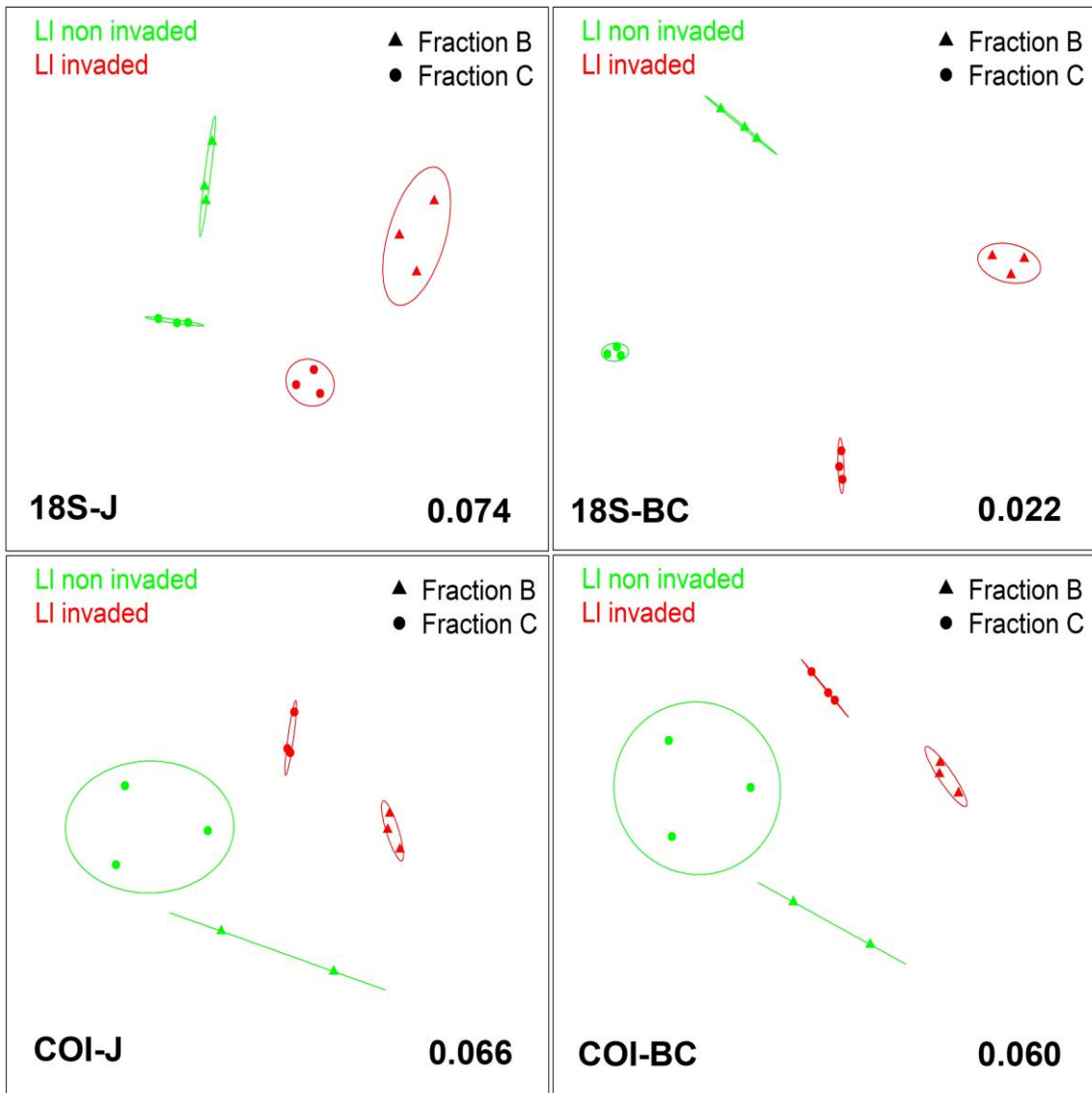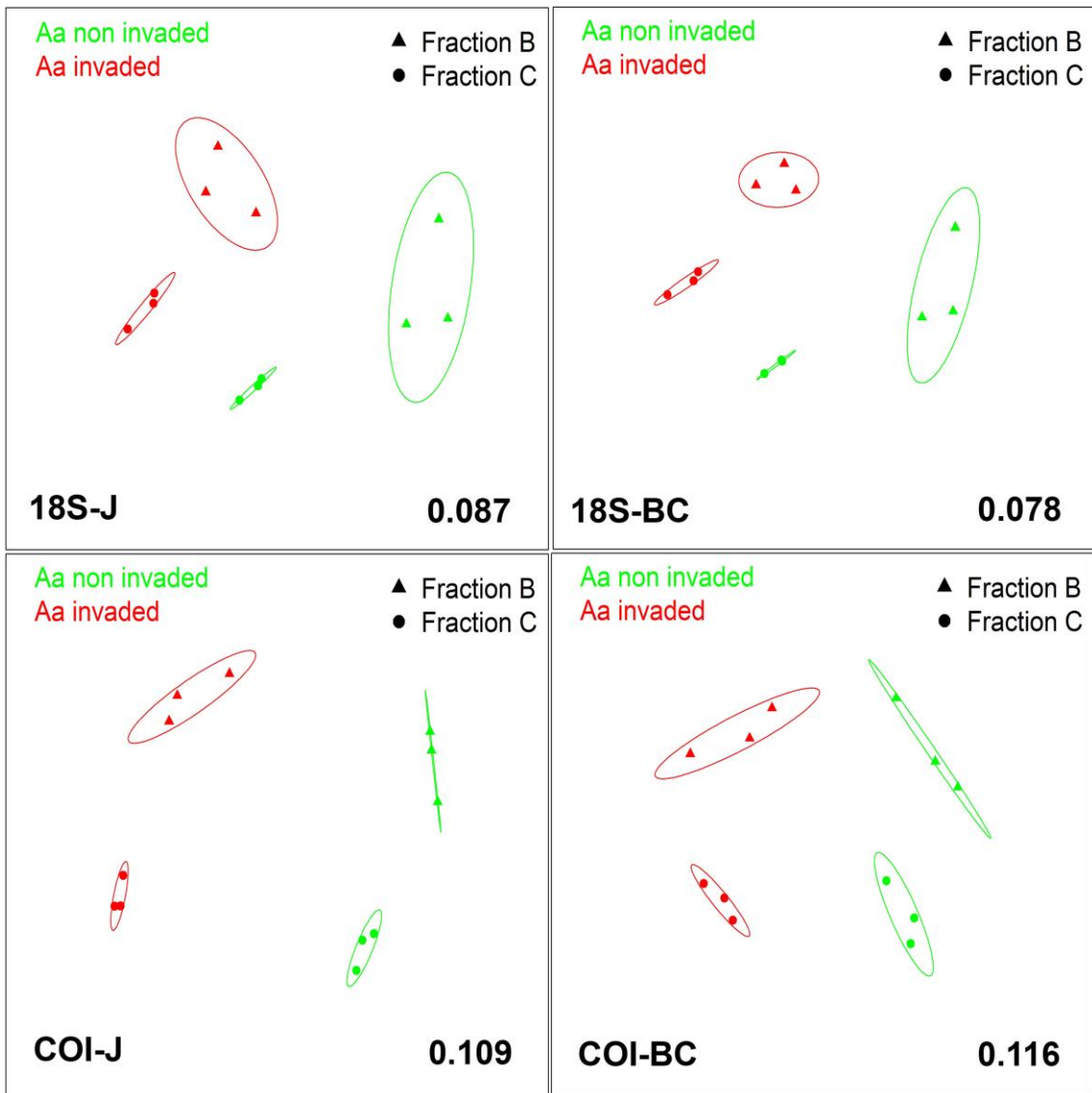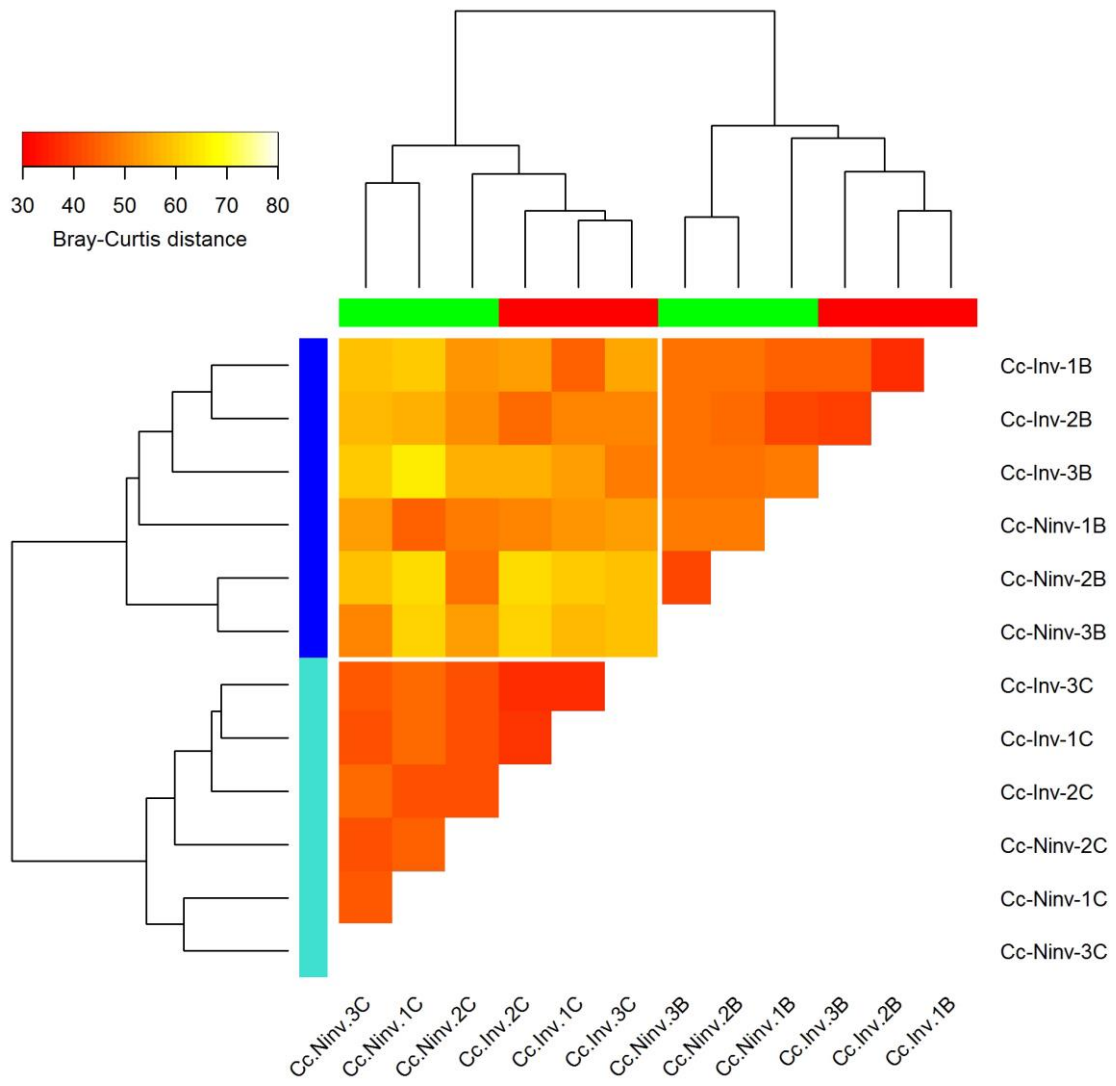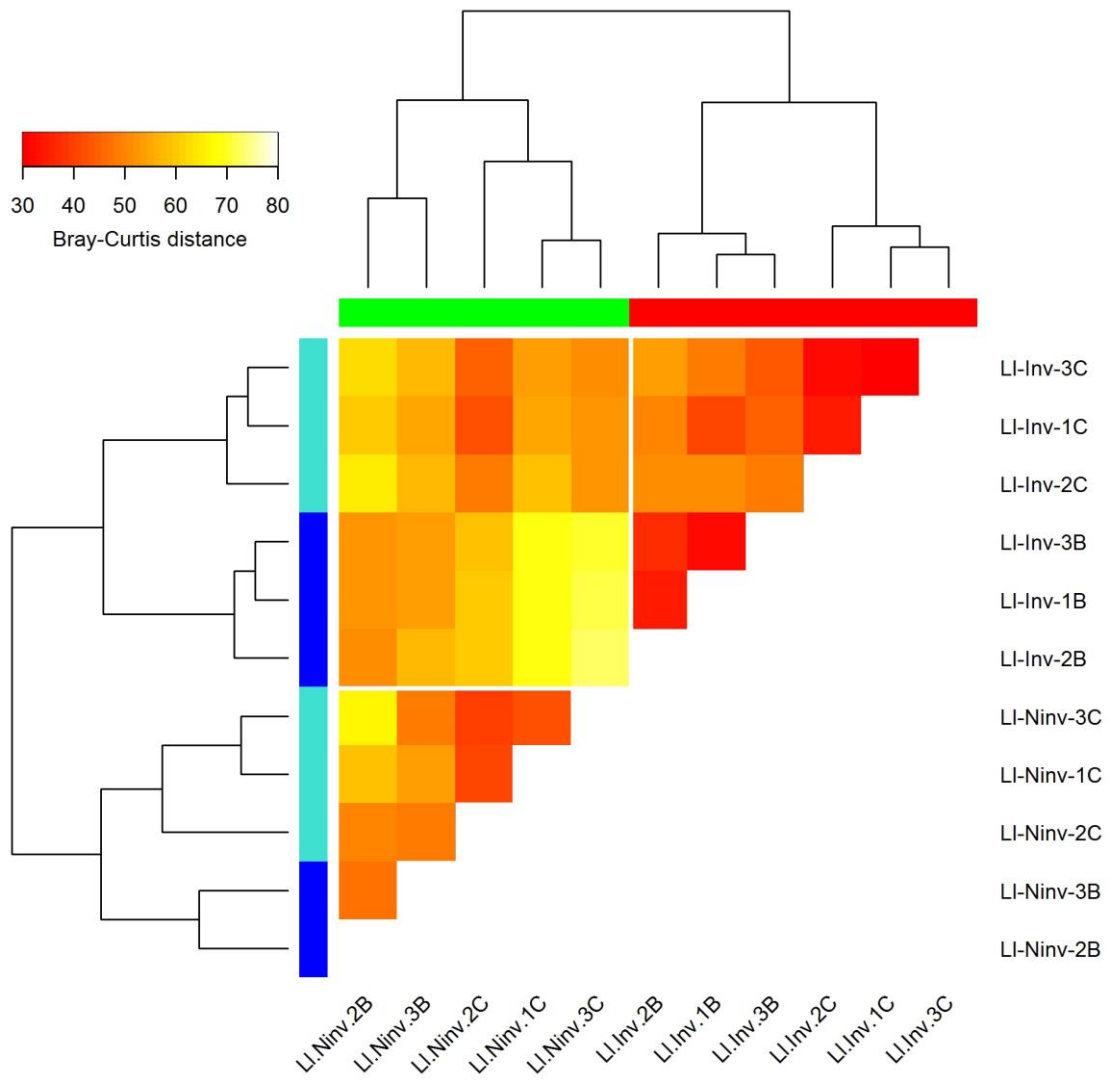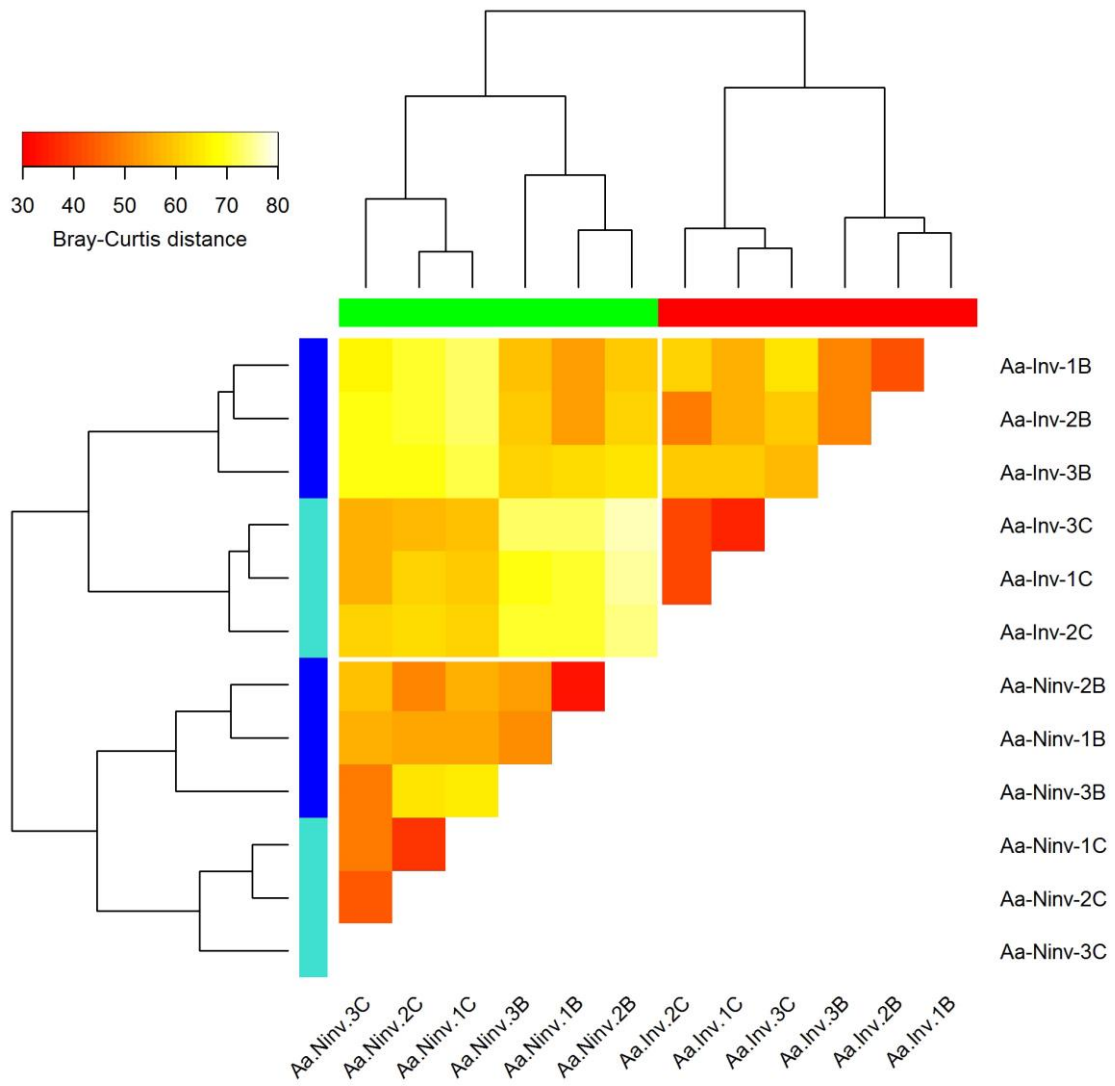
FIGURE 1

FIGURE 2

FIGURE 3
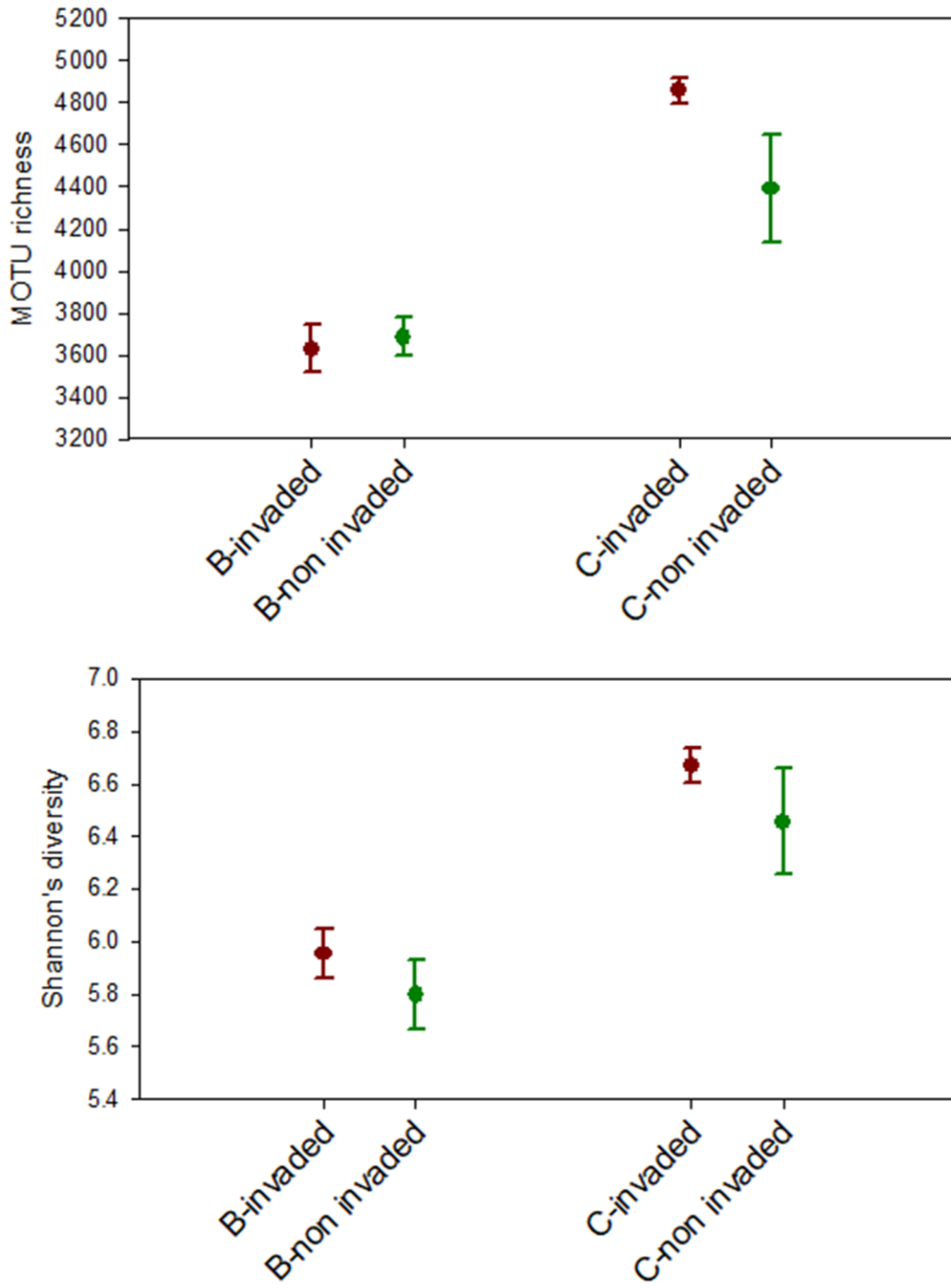
FIGURE 4

FIGURE 5

FIGURE 6

Caulerpa cylindracea

FIGURE 7

Lophocladia lallemandii

FIGURE 8

Asparagopsis armata

FIGURE 9

FIGURE S1

FIGURE S2

Table S3. PERMANOVA tests of the factors Condition (invaded or non-invaded, with Replicate as nested blocking factor) and Fraction (B or C) for the Jaccard similarity on presence-absence data based on 18S. PERMDISP tests of main factors are also included.

Cc dataset	df	SS	Pseudo-F	P-value	Permdisp
Condition	1	2714.6	1.641	0.181	0.368
Fraction	1	5466.4	5.319	0.012	0.004
Replicate (Condition)	4	6617.3	1.610	0.177	
Condition*Fraction	1	1182.9	1.151	0.353	
Residual	4	4110.7			

Ll dataset	df	SS	Pseudo-F	P-value	Permdisp
Condition	1	5934.6	3.541	0.009	0.488
Fraction	1	5120.0	4.588	0.016	0.004
Replicate (Condition)	4	1675.9	1.502	0.199	
Condition*Fraction	1	1728.5	1.549	0.250	
Residual	4	1115.9			

Aa dataset	df	SS	Pseudo-F	P-value	Permdisp
Condition	1	5243.8	2.890	0.020	0.638
Fraction	1	5103.7	4.348	0.013	0.004
Replicate (Condition)	4	7257.3	1.546	0.193	
Condition*Fraction	1	1800.8	1.534	0.228	
Residual	4	4695.5			

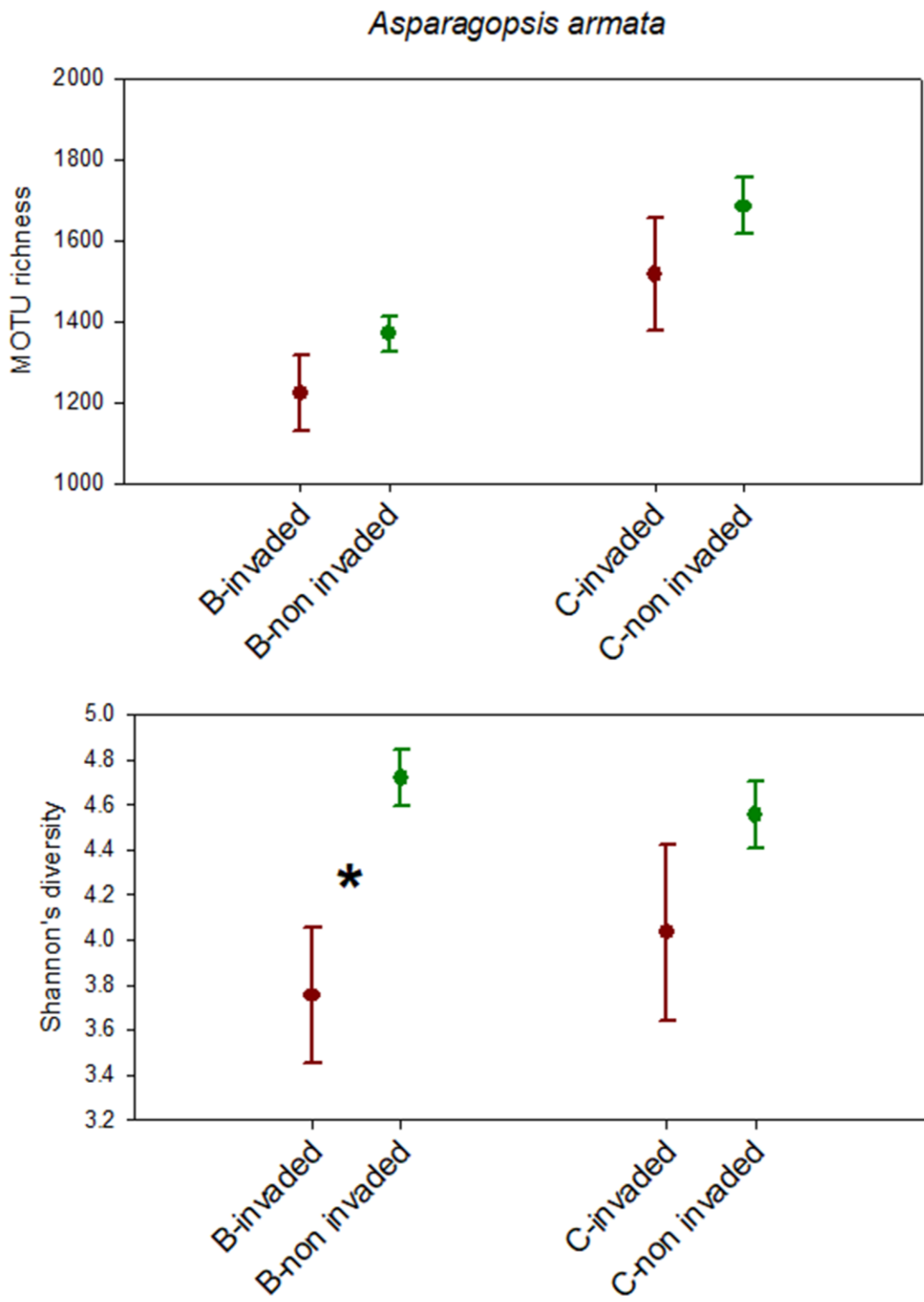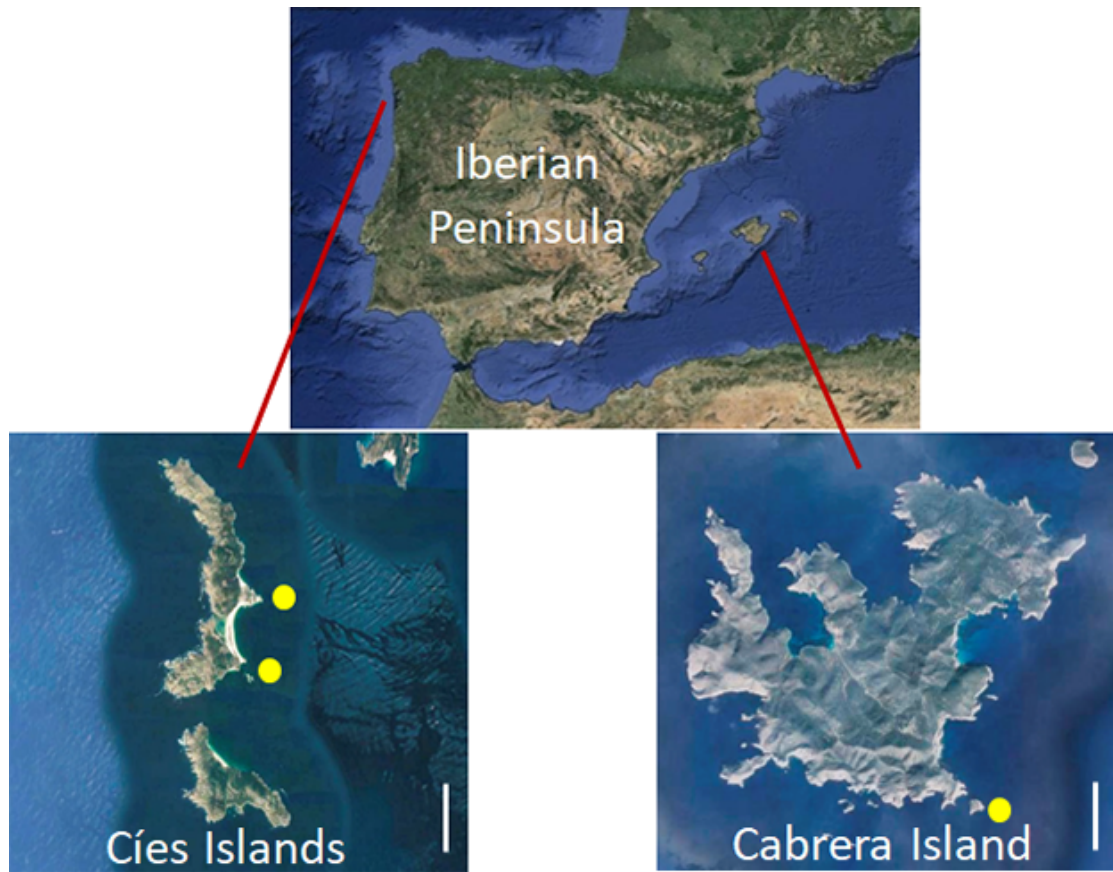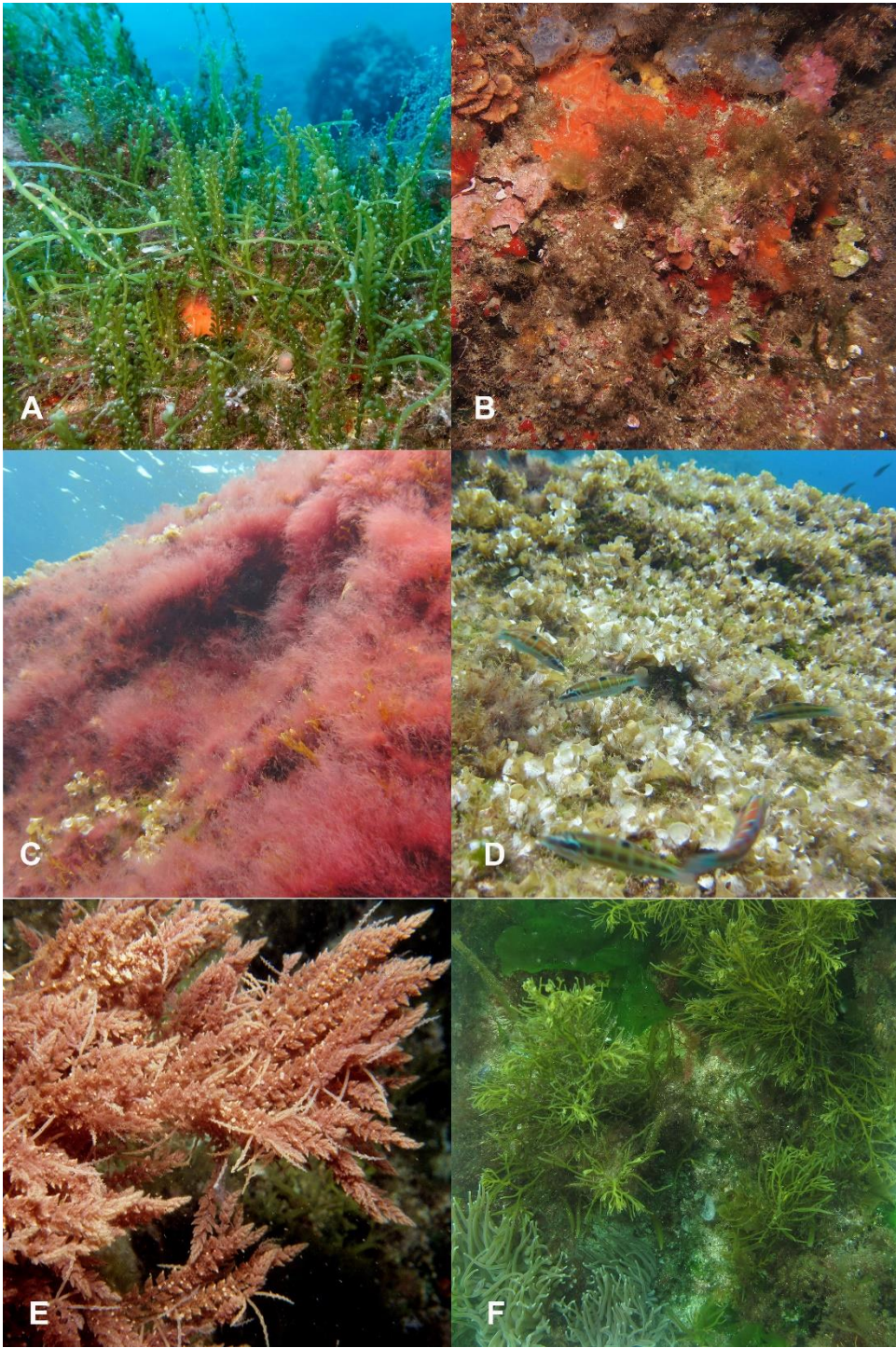Table S4. PERMANOVA tests of the factors Condition (invaded or non-invaded, with Replicate as nested blocking factor) and Fraction (B or C) for the Jaccard similarity on presence-absence data based on COI. PERMDISP tests of main factors are also included.

Cc dataset	df	SS	Pseudo-F	P-value	Permdisp
Condition	1	2714.6	1.641	0.181	0.368
Fraction	1	5466.4	5.319	0.012	0.004
Replicate (Condition)	4	6617.3	1.610	0.177	
Condition*Fraction	1	1182.9	1.151	0.353	
Residual	4	4110.7			

Ll dataset	df	SS	Pseudo-F	P-value	Permdisp
Condition	1	4494.1	2.550	0.033	0.157
Fraction	1	4368.6	3.589	0.072	0.985
Replicate (Condition)	4	7253.2	1.490	0.295	
Condition*Fraction	1	1544.6	1.269	0.347	
Residual	3	3652.0			

Aa dataset	df	SS	Pseudo-F	P-value	Permdisp
Condition	1	6115.9	2.915	0.016	0.334
Fraction	1	5984.3	4.147	0.018	0.220
Replicate (Condition)	4	8392.4	1.454	0.218	
Condition*Fraction	1	1816.5	1.259	0.326	
Residual	4	5772.8			

Table S5. List of the 25 most abundant (in reads) MOTUs among the ones identified by SIMPER as contributing more to the differentiation between the invaded and non-invaded communities of the Cc dataset using the COI gene. Id is the identifier of the MOTU as in Table S2. Abundance 1 and 2 are the mean abundances (in semiquantitative ranks) of the MOTUs in the non-invaded (1) and invaded (2) communites. Total reads in the Cc dataset, and the identity of the MOTU with the best-match entry in the reference database are indicated. The super-group and phylum assigned to each MOTU is also given. Rank refers to the taxonomic rank that ecotag could assign to the MOTUs, taxid is the identifier (as in NCBI taxonomy browser) of the taxon assigned to the MOTUs, and name is the corresponding name given in the NCBI taxonomy browser.

id	abundance 1	abundance 2	reads Cc dataset	best_identity	Supergroup	Phylum	rank	taxid	name
MBPA_001421159	3.33	0.67	8665	0.997	Archaeplastida	Rhodophyta	species	348032	*Mesophyllum macroblastum*
MBPA_003786865	0	3.83	7138	0.997	Metazoa	Porifera	species	3722	*Crambe crambe*
MBPA_001709134	1	4	5913	0.981	Metazoa	Cnidaria	species	252667	*Clytia linearis*
MBPA_000444767	3.17	1.33	4232	0.864	Stramenopiles	Ochrophyta	Class	2870	Phaeophyceae
MBPA_005146432	0.67	2.67	3691	0.997	Metazoa	Mollusca	species	71494	*Elysia viridis*
MBPA_004493206	1.17	3.33	2894	0.822	Metazoa	Annelida	Class	6341	Polychaeta
MBPA_000264855	0.67	3.83	2793	0.813	Metazoa	Bryozoa	Phylum	6072	Bryozoa
MBPA_002804915	0.67	3.33	2712	0.810	Metazoa	Annelida	Phylum	6340	Annelida
MBPA_000006344	1.33	3.83	2514	0.876	Metazoa	Arthropoda	Order	6683	Decapoda
MBPA_003826858	0.33	3.67	2314	0.997	Archaeplastida	Rhodophyta	species	173553	*Womersleyella setacea*
MBPA_000035239	0.67	2.83	2011	0.899	Metazoa	Nemertea	Order	6227	Monostilifera
MBPA_000219187	3.17	0.5	1677	0.789	Metazoa	NA	no rank	33317	Protostomia
MBPA_006025192	0.67	3.33	1624	0.751	Metazoa	Bryozoa	Order	97265	Cyclostomatida
MBPA_000876839	0.67	3.17	1539	0.826	Metazoa	Bryozoa	Order	10207	Cheilostomatida
MBPA_000409671	0.5	3.67	1517	1.000	Metazoa	Echinodermata	species	70180	*Ophiothrix fragilis*
MBPA_001534681	0.33	3	1478	0.997	Metazoa	Annelida	species	188455	*Lysidice ninetta*
MBPA_002466875	1	3.83	1415	0.867	Metazoa	NA	no rank	6072	Eumetazoa
MBPA_002241983	2.67	0	1384	1.000	Metazoa	Arthropoda	species	10000027	*Dexamine spiniventris*
MBPA_002453200	0.67	3.83	1235	0.904	Metazoa	Cnidaria	Order	500008	Leptothecata
MBPA_000145110	3.33	0	1190	0.793	Metazoa	NA	no rank	33317	Protostomia
MBPA_006021345	2.67	0	1172	0.852	Metazoa	Annelida	species	1311521	*Polycirrus carolinensis*
MBPA_000464416	0	2.67	922	0.833	Metazoa	Cnidaria	Phylum	33208	Cnidaria
MBPA_000001158	3.33	1.17	892	1.000	Metazoa	Bryozoa	Order	10207	Cheilostomatida
MBPA_004323452	2.83	0.67	889	0.858	Metazoa	Cnidaria	Subclass	37516	Hydroidolina
MBPA_005851436	0	2.83	860	0.830	Metazoa	NA	no rank	33317	Protostomia

Table S6. List of the 25 most abundant (in reads) MOTUs among the ones identified by SIMPER as contributing more to the differentiation between the invaded and non-invaded communities of the LI dataset using the COI gene. Id is the identifier of the MOTU as in Table S2. Abundance 1 and 2 are the mean abundances (in semiquantitative ranks) of the MOTUs in the non-invaded (1) and invaded (2) communites. Total reads in the LI dataset, and the identity of the MOTU with the best-match entry in the reference database are indicated. The super-group and phylum assigned to each MOTU is also given. Rank refers to the taxonomic rank that ecotag could assign to the MOTUs, taxid is the identifier (as in NCBI taxonomy browser) of the taxon assigned to the MOTUs, and name is the corresponding name given in the NCBI taxonomy browser.

id	abundance 1	abundance 2	reads LI dataset	best_identity	Supergroup	Phylum	rank	taxid	name
MBPA_000013571	0	2.67	7653	0.855	Metazoa	Cnidaria	Phylum	6073	Cnidaria
MBPA_006189180	0	2.67	6605	0.997	Archaeplastida	Rhodophyta	species	389204	*Peyssonnelia rubra*
MBPA_007976277	0	3.33	4200	0.997	Metazoa	Cnidaria	species	1267107	*Cornularia cornucopiae*
MBPA_001991748	1.4	3.83	3428	0.826	Metazoa	Bryozoa	Order	10207	Cheilostomatida
MBPA_005434469	3.8	0.67	2883	0.981	Metazoa	Porifera	Genus	121493	*Oscarella*
MBPA_003778014	4	0	2570	0.997	Stramenopiles	Ochrophyta	species	10000052	*Cystoseira* sp.2
MBPA_000660468	0.8	3.33	2341	0.907	Archaeplastida	Rhodophyta	Order	31395	Corallinales
MBPA_000528756	1.6	3.83	2286	0.870	Metazoa	Cnidaria	Subclass	37516	Hydroidolina
MBPA_000372200	3	0	1994	0.911	Archaeplastida	Rhodophyta	Order	28017	Gigartinales
MBPA_003956246	0.8	3.5	1492	0.895	Metazoa	Cnidaria	Order	500008	Leptothecata
MBPA_004237361	1.2	4	1297	0.968	Metazoa	Cnidaria	Genus	1427710	Parasphaerasclera
MBPA_008110653	1.4	3.83	1264	0.876	Archaeplastida	Rhodophyta	Order	2802	Ceramiales
MBPA_008290877	0	3.33	1264	0.846	Metazoa	Cnidaria	Subclass	37516	Hydroidolina
MBPA_004139193	1.2	4	1095	0.971	Metazoa	Cnidaria	Genus	1427710	*Parasphaerasclera*
MBPA_003796020	0	3.83	1061	0.997	Archaeplastida	Rhodophyta	species	10000010	*Rhodophyllis* sp.MAR3
MBPA_000025963	2.4	0	1035	1.000	Archaeplastida	Rhodophyta	species	1347097	*Phymatolithon* sp. 1CPVP
MBPA_001107383	1	3.83	1023	0.971	Metazoa	Cnidaria	Genus	1427710	*Parasphaerasclera*
MBPA_008481052	0.8	3.67	1023	0.892	Metazoa	Cnidaria	Order	500008	Leptothecata
MBPA_000149612	4	0	1015	0.994	Alveolata	Myzozoa	species	671363	*Symbiodinium* sp. clade A3
MBPA_006682039	4	0	1001	0.892	Metazoa	Cnidaria	species	498518	*Clava multicornis*
MBPA_000058511	1	4	952	0.898	Archaeplastida	Rhodophyta	Order	31395	Corallinales
MBPA_000268661	1.4	3.83	946	0.876	Archaeplastida	Rhodophyta	Order	2802	Ceramiales
MBPA_004787675	1.6	4	928	0.880	Archaeplastida	Rhodophyta	Order	2802	Ceramiales
MBPA_005744111	1	3.83	918	0.968	Metazoa	Cnidaria	Genus	1427710	*Parasphaerasclera*
MBPA_005129138	1.2	3.83	872	0.971	Metazoa	Cnidaria	Genus	1427710	*Parasphaerasclera*

Table S7. List of the 25 most abundant (in reads) MOTUs among the ones identified by SIMPER as contributing more to the differentiation between the invaded and non-invaded communities of the Aa dataset using the COI gene. Id is the identifier of the MOTU as in Table S2. Abundance 1 and 2 are the mean abundances (in semiquantitative ranks) of the MOTUs in the non-invaded (1) and invaded (2) communites. Total reads in the Aa dataset, and the identity of the MOTU with the best-match entry in the reference database are indicated. The super-group and phylum assigned to each MOTU is also given. Rank refers to the taxonomic rank that ecotag could assign to the MOTUs, taxid is the identifier (as in NCBI taxonomy browser) of the taxon assigned to the MOTUs, and name is the corresponding name given in the NCBI taxonomy browser.

id	abundance 1	abundance 2	reads Aa dataset	best_identity	Supergroup	Phylum	rank	taxid	scientific_name
MBPA_004522023	4	1.33	35946	0.997	Metazoa	Arthropoda	species	10000084	*Dynamene magnitorata*
MBPA_000294452	3.33	0	13329	0.905	Metazoa	Cnidaria	Class	6072	Hydrozoa
MBPA_006987114	0.67	3.17	9887	0.997	Metazoa	Mollusca	species	256138	*Callochiton septemvalvis*
MBPA_004445160	4	0	8858	0.997	Metazoa	Arthropoda	species	483407	*Pilumnus hirtellus*
MBPA_003258648	4	1.33	8375	0.997	Metazoa	Chordata	species	1440460	*Morchellium argus*
MBPA_004437216	3.33	0	8235	0.997	Metazoa	Arthropoda	species	371503	*Pseudoprotella phasma*
MBPA_000029197	3.33	0.67	6541	0.836	Metazoa	Arthropoda	Order	29979	Isopoda
MBPA_000001477	1.67	4	4708	0.997	Archaeplastida	Rhodophyta	species	31364	*Bonnemaisonia hamifera*
MBPA_002899382	2.67	0	3490	0.997	Metazoa	Annelida	species	251482	*Sabellaria alveolata*
MBPA_002374119	0.83	3.33	3280	1.000	Metazoa	Echinodermata	species	7675	*Psammechinus miliaris*
MBPA_003823224	0.83	3.67	3278	0.873	Archaeplastida	Rhodophyta	Class	2806	Florideophyceae
MBPA_005781195	0.67	3.17	3081	0.838	Metazoa	Arthropoda	Order	6821	Amphipoda
MBPA_000318927	3.33	0	3020	0.882	Metazoa	Cnidaria	species	308574	*Plumularia setacea*
MBPA_000555161	0.5	3.83	2478	0.997	Metazoa	Chordata	species	581068	*Distomus variolosus*
MBPA_000061739	3.83	0	2266	0.997	Metazoa	Arthropoda	species	371449	*Jassa falcata*
MBPA_003881323	1.33	3.67	2192	0.984	Metazoa	Arthropoda	order	116167	Sessilia
MBPA_001617979	3.83	1.33	1949	0.754	Metazoa	Mollusca	Class	6544	Bivalvia
MBPA_000274757	3.67	0	1896	0.994	Metazoa	Mollusca	species	6551	*Mytilus trossulus*
MBPA_000012103	0.83	3.33	1752	0.851	Metazoa	Arthropoda	Order	6821	Amphipoda
MBPA_002454198	1	3.17	1583	0.842	Metazoa	Arthropoda	Order	6821	Amphipoda
MBPA_003301051	0.5	3.67	1370	0.839	Metazoa	Arthropoda	Order	6821	Amphipoda
MBPA_000016104	1.17	3.83	1349	0.997	Archaeplastida	Rhodophyta	species	2802	*Dasya tenuous*
MBPA_003761857	0	3.33	1316	0.882	Metazoa	Arthropoda	Order	75398	Tanaidacea
MBPA_000763146	0	4	1258	0.851	Metazoa	Cnidaria	Class	6072	Hydrozoa
MBPA_000149612	4	0.5	1173	0.994	Alveolata	Myzozoa	species	671363	*Symbiodinium* sp. clade A3