Pan-regional marine benthic cryptobiome biodiversity patterns revealed by metabarcoding Autonomous Reef Monitoring Structures

J.K. Pearman, G. Chust, E. Aylagas, E. Villarino, J.R. Watson, A. Chenuil, A. Borja, A.E. Cahill, L. Carugati, R. Danovaro, et al.

DR. JOHN KENNETH PEARMAN (Orcid ID : 0000-0002-2237-9723)
DR. EVA AYLAGAS (Orcid ID : 0000-0001-9792-8451)
DR. NAIARA RODRIGUEZ-EZPELETA (Orcid ID : 0000-0001-6735-6755)

Article type : Original Article

Pan-regional marine benthic cryptobiome biodiversity patterns revealed by metabarcoding Autonomous Reef Monitoring Structures

Running Title: Pan-regional ARMS biodiversity

Pearman, J.K.[1,2], Chust, G.[3], Aylagas, E.[1], Villarino, E.[3,9,10], Watson, J.R.[9], Chenuil, A.[4], Borja, A.[5], Cahill, A.E.[6], Carugati, L.[7], Danovaro, R.[7], David, R.[4], Irigoien, X.[3,5], Mendibil, I.[3], Moncheva, S.[8], Rodríguez-Ezpeleta, N. [3], Uyarra, M.C.[5], Carvalho, S.[1]

1 King Abdullah University of Science and Technology (KAUST), Red Sea Research Center, Thuwal, 23955-6900 Saudi Arabia
2 Coastal and Freshwater Group, Cawthron Institute, Private Bag 2, Nelson 7042, New Zealand
3 AZTI, Basque Research and Technology Alliance (BRTA) - Marine Research, Herrera Kaia, Portualdea z/g – 20110 Pasaia (Gipuzkoa), Spain
4 Institut Méditerranéen de Biodiversité et d'Ecologie marine et continentale (IMBE), Aix Marseille Univ, Avignon Université, CNRS, IRD, IMBE, Marseille, France
5 IKERBASQUE, Basque Foundation for Science, Bilbao, Spain
6 Biology Department, Albion College, Albion, Michigan, USA
7 Dipartimento di Scienze della Vita e dell'Ambiente, Università Politecnica delle Marche, Ancona, Italy

8 Institute of Oceanology, (IO-BAS), First May Street 40, P.O.Box 152, Varna 9000, *Bulgaria*
9 College of Earth, Ocean and Atmospheric Sciences, Oregon State University, Corvallis, OR, USA, 97330
10 *Scripps Institution of Oceanography & Section of Ecology, Behavior and Evolution University of California San Diego, 9500 Gilman Drive #0218 | La Jolla, CA 92093*

***Corresponding Author*:** Susana Carvalho, King Abdullah University of Science and Technology, Red Sea Research Center, Thuwal 23955-6900, Kingdom of Saudi Arabia
Telephone +966-(12) 8082908
Email: susana.carvalho@kaust.edu.sa

***Keywords*:** biodiversity; ecosystem connectivity; climate change; metapopulation; dispersal limitation; ecological niche; ecological monitoring; mitochondrial cytochrome oxidase I (COI)

Abstract:

Autonomous Reef Monitoring Structures (ARMS) have been applied worldwide to characterize the critical yet frequently overlooked biodiversity patterns of marine benthic organisms. In order to disentangle the relevance of environmental factors in benthic patterns, here, through standardized metabarcoding protocols, we analyze sessile and mobile (<2 mm) organisms collected using ARMS deployed across six regions with different environmental conditions (3 sites x 3 replicates per region): Baltic, Western Mediterranean, Adriatic, Black and Red Seas, and the Bay of Biscay. A total of 27473 Amplicon Sequence Variants (ASVs) were observed ranging from 1404 in the Black Sea to 9958 in the Red Sea. No ASVs were shared amongst all regions. The highest number of shared ASVs was between the Western Mediterranean and the Adriatic Sea (116) and Bay of Biscay (115). Relatively high numbers of ASVs (103), mostly associated with the genus *Amphibalanus*, were also shared between the lower salinity seas (Baltic and Black Seas). We found that compositional differences in spatial patterns of rocky-shore benthos are

determined slightly more by dispersal limitation than environmental filtering. Dispersal limitation was similar between sessile and mobile groups, while the sessile group had a larger environmental niche breadth than the mobile group. Further, our study can provide a foundation for future evaluations of biodiversity patterns in the cryptobiome, which can contribute up to 70% of the local biodiversity.

Introduction

Oceanic and coastal areas are essential sources of goods and services for human well-being (Barbier, Hacker, Koch, Stier, & Silliman, 2011; Costanza et al., 1997), but are also affected by human pressures (Halpern et al., 2008; Korpinen & Andersen, 2016; Lotze, Guest, O'Leary, Tuda, & Wallace, 2018). These pressures highlight the need for scientifically informed conservation and management efforts (Costello & Wilson, 2011). Understanding what shapes biodiversity is vital so that changes in the status of biological communities can be anticipated and managed (Andersen, Halpern, Korpinen, Murray, & Reker, 2015; Costello & Wilson, 2011; Elliott, 2014; Micheli et al., 2013), for example through the protection of species on the brink of extinction (Costello & Wilson, 2011). Contrary to terrestrial domains, biogeographic barriers are relatively limited in the oceans making marine ecosystems particularly vulnerable to the effects of local disasters, for example oil spills and fisheries overharvesting, expanding over large distances (Cordes et al., 2016; Sammarco et al., 2013). Thus, the development of data-driven and standardized environmental monitoring tools, to maintain natural levels of biodiversity within nearshore ecosystems, is of paramount importance (Danovaro et al., 2016).

Traditional monitoring techniques in hard-bottom marine environments have mainly been based on visual census and morphological identification of the most conspicuous organisms present (e.g. macroalgae, corals, sponges, fish) along transects (that vary in length, width and also on the method - e.g. photo-transects *versus* line intercept method) (Danovaro et al., 2016). However, a high proportion of the benthic biodiversity in these systems comprises small sessile, encrusting or mobile organisms (Enochs, Toth, Brandtneris, Afflerbach, & Manzello, 2011). These organisms are considered to be part of the cryptobiome (Carvalho et al., 2019) as they inhabit cavities

(cryptic habitats) within the rocky architecture for temporary shelter (e.g. nocturnal species) or as a source of food, and are often neglected during traditional surveys (Pearman, Anlauf, Irigoien, & Carvalho, 2016; Reaka-Kudla, 1997). The cryptobiome encompasses a diverse selection of ecologically important groups such as suspension feeders (Richter, Wunsch, Rasheed, Kötter, & Badran, 2001; Scheffers, de Goeij, van Duyl, & Bak, 2003), predators (Reaka, 1987), herbivores (Coen, 1988) and detritivores (Rothans & Miller, 1991). Because of their diverse ecological roles, small size and fast generation times (Finlay, 2002), the responses of the cryptobiome to environmental stressors may differ from those of the larger macro-organisms usually studied. Despite their importance in benthic ecosystems, the cryptic nature of being small and difficult to spot, as well as the diversity of the phyla represented, requires specialized taxonomists for identification, creating a bottleneck which can limit both the temporal and spatial scales at which studies can be undertaken.

Recognizing the urgency for standardized methods to comprehensively assess hard bottom benthic biodiversity across different habitats and regions, the Coral Reef Ecosystem Division (CRED) of the United States National Oceanic and Atmospheric Administration (NOAA) developed the Autonomous Reef Monitoring Structures (ARMS; https://www.pifsc.noaa.gov/cred/survey_methods/arms/overview.php). The alternating open and obstructed format in the gaps between plates comprising the ARMS was designed to mimic the structural complexity of hard-bottom substrata allowing the colonization of a variety of organisms with different niche preferences (Zimmerman & Martin, 2004). ARMS-associated communities can be analyzed either morphologically (David et al., 2019) or through DNA metabarcoding (e.g. targeting a short DNA fragment of the mitochondrial cytochrome oxidase I (COI) gene) to identify the whole spectrum of their biodiversity and community composition of sessile and mobile organisms (Leray & Knowlton, 2015). In addition, with the deployment of ARMS across large spatial scales and the use of standardized sampling protocols, large scale biodiversity patterns can be obtained. So far this has been undertaken using molecular approaches along the length of the Red Sea (Carvalho et al., 2019; Pearman et al., 2019) with morphological assessments of benthic substrates being undertaken in European waters and the Red Sea (David et al., 2019). ARMS have been deployed globally (https://www.oceanarms.org/deployments/search, accessed 07/07/2020) opening the possibility for global studies. Comparisons between ARMS and coral dead head communities show the same average similarity as comparisons between dead head communities

(Plaisance, Caley, Brainard, & Knowlton, 2011) while similar microphytobenthos communities have been observed on ARMS as found in natural substrates (Pennesi & Danovaro, 2017). ARMS have thus been increasingly applied over the last decade to research coral habitats in different regions (e.g. east coast of the USA (Leray & Knowlton, 2015); Red Sea (Carvalho et al., 2019; Pearman et al., 2016, 2019); French Polynesia (Ransome et al., 2017); Gulf of Aqaba (Al-Rshaidat et al., 2016), Indonesia (Hazeri et al., 2019) and European waters (David et al., 2019; Pennesi & Danovaro, 2017)). The combination of a standardized tool for sampling biodiversity (i.e. ARMS) and metabarcoding-based characterization of the benthic biota has great potential to allow for assessments of biodiversity to be more spatially and temporally comprehensive, allowing to tackle questions at both local and global scales.

The ARMS monitoring system provides a standardized methodology to disentangle the relative importance of dispersal limitation (Wilson & MacArthur, 1967) and niche (Hutchinson, 1957) adaptive processes to the assembly of biological communities (Burgess, Baskett, Grosberg, Morgan, & Strathmann, 2016; Chave, 2004). Dispersal plays a key role in connecting populations, in maintaining species coexistence, and promoting regional biodiversity (Mouquet & Loreau, 2003). However, the dispersal strategies of many marine species are still unknown due to the difficulty of tracking the trajectory and fate of propagules (Selkoe & Toonen, 2011; Weersing & Toonen, 2009). According to the neutral theory of biodiversity, dispersal limitation of communities can be inferred from the logarithmic decline of species composition similarity among sites (i.e., the opposite of beta diversity) with increasing geographical distance (termed "distance decay"), when migration rate is low (Hubbell, 1997, 2001). In agreement with this view, previous studies have shown that dispersal-related traits of species (e.g. sessile *vs* mobile, pelagic larval duration, size) affect their biogeographic distribution (Kinlan & Gaines, 2003), and this distribution can be indirectly estimated at community level (Chust et al., 2016; Villarino et al., 2018). In contrast, the ecological niche theory (Pocheville, 2015) assumes that differences in species composition among communities is caused by heterogeneity in the environment or limiting resources, and by environmental filtering of species according to their environmental requirements, such as climate, oceanographic and coastal conditions, and competition for resources such as nutrients for marine algae. Moreover, historical events such as glacial periods have been shown to affect taxonomic patterns over large spatial scales (Bestová, Munoz, Svoboda, Škaloud, & Violle, 2018; Normand et al., 2011). The regional distribution of species arises from

limitations to dispersal or niche adaptive processes and has been a long-standing and challenging debate in ecology as space is often correlated with environment (Hubbell, 2001; Legendre & Legendre, 2012). The relative importance of the two processes shaping the community distribution and species turnover can differ according to the community, scale of study, or habitat (terrestrial or marine); yet, such studies in hard-bottom communities are scarce.

Here we investigate the composition and spatial species turnover (beta diversity) of the cryptobiome, comprising sessile and mobile organisms (<2 mm) which inhabit hard-bottoms. We used standardized methods to address those communities, specifically ARMS combined with a metabarcoding approach targeting the mitochondrial cytochrome oxidase I (COI) gene. The study was undertaken on a pan-regional scale ranging from the Baltic Sea (~56°N) to the Red Sea (~21°N) and from the Bay of Biscay (NE Atlantic, ~2°W) to the Black Sea (~28°E), including also Western Mediterranean and Adriatic Seas. We tested whether environmental conditions are more important than dispersal limitation in explaining community composition, and whether the relative contribution of both components (dispersal and environmental) is group (sessile and mobile) dependent. The pan-regional domain of this study enables a unique opportunity to investigate benthic biodiversity patterns across a broad range of environmental gradients and hierarchical spatial scales using a standardized approach, and to test which fundamental ecological processes are responsible for setting observed patterns in biodiversity of an overlooked biological component of hard bottom environments.

Methods

ARMS units comprised of nine 22.5 x 22.5 PVC plates stacked on top of a 35 X 45 cm base plate with spacers separating the plates. Bars were placed on each alternate level to create closed different flow regimes. The construction of the ARMS was as described on the Smithsonian website (https://www.oceanarms.org/protocols/arms-assembly; accessed 07/07/2020). ARMS units were deployed in 6 regions. Within each region deployment was undertaken at 3 locations separated by around 10-20 km, with 3 replicates at each location, separated by 5-10 m; ARMS were deployed at a depth ranging from 7 to 19 m (Figure 1 and Supplementary Table 1), generally for 12-14 months between May 2013 and June 2014 (see Supplementary Table 1 for details at each site including the predominant substrate the ARMS were deployed on). Due to exceptional weather conditions, recovery of the Baltic Sea samples was performed after a submersion period

of 26 months. Storms in the Adriatic caused the loss of the deployed ARMS and thus another set were deployed the following year. ARMS were deployed and retrieved by scuba. During retrieval the ARMS unit was covered so that mobile organisms within the ARMS could not escape. Upon retrieval the ARMS were placed in a large container filled with 106 µm filtered seawater (from the site) for transport back to the laboratory. ARMS were processed by representative institutes within each region and processed following the protocol described in Leray and Knowlton, (2015). Visual analysis of all the ARMS showed them to be extensively colonized, the extent of which has been described for a subset of these ARMS by David et al., (2019). Briefly, the ARMS were dissembled within water obtained from the sampling site and the plates submerged in 0.2 µm filtered seawater. The plates were gently brushed to remove any mobile organisms. These mobile organisms were retained to be sieved. Sessile organisms growing on the plates were scraped into a tray and homogenized using a blender with approximately 40 g of tissue, being preserved in DMSO buffer. The water containing the ARMS was filtered through a stack of sieves (2mm, 500 µm and 106 µm). The largest fraction was individually photographed and stored in 95% ethanol, but is not considered in this study. The small fractions were bulk preserved in 95% ethanol before being separated from sediments by decantation. The organic fraction was crushed using a mortar and pestle. Environmental DNA was extracted from 10 g of the sessile fraction (comprising the homogenized bulk sample scraped from the plates) and from two size fractions of crushed bulk mobile organisms (106-500 µm and 500-2000 µm) using the Powermax Soil DNA kit (MO BIO), as per the manufacturer's instructions with the exception of the bead-beating step, which was replaced by shaking incubation overnight at 56 °C with the addition of Proteinase K (0.4 mg/mL). DNA was purified using the Powerclean DNA Clean-up kit (MO BIO) and quantified with a Quibit Fluorometer (Invitrogen). A versatile primer set amplifying a 313 bp fragment of the COI gene (Forward: GGWACWGGWTGAACWGTWTAYCCYCC; Reverse: TAIACYTCIGGRTGICCRAARAAYCA; (Geller, Meyer, Parker, & Hawk, 2013; Leray et al., 2013). The PCR profile consisted of an initial 3 min denaturation step at 98 °C, followed by 27 cycles at 98 °C for 10 sec, 46 °C for 45 sec and, 72 °C for 45 sec, with a final extension of 5 min at 72 °C. All PCR reactions were done in triplicate, using 0.4 µl of 10 µM primers, 10 µl of Phusion High Fidelity Mix (2X), 7.2 µl of water and 2 µl (~ 10 ng) of DNA. PCR triplicates were combined and cleaned up using AMPure beads. A second PCR was undertaken following the Illumina 16S metagenomics library prep protocol to attach Nextera XT Illumina multiplex tags or in the case of the Adriatic Sea sample the Ovation Rapid DR Multiplex System (NuGen)

(Supplementary Table 2). Samples were normalized and pooled before samples were diluted to 4 nM and denatured as per the Illumina 16S metagenomics library prep manual. A 10-20% PhiX spike was added to the library. Sequencing was undertaken on an Illumina MiSeq at the King Abdullah University of Science and Technology (KAUST) Bioscience Core Laboratory, Centre d'Innovation Génome Québec, McGill University, LGC Genomics GmbH (Berlin, Germany) and SGiker Genotyping and Sequencing Unit (EHU/UPV, ERDF and ESF). Sequences used for this study are deposited in the NCBI short read archive under accession PRJNA557002.

For each sequencing run sequences were automatically demultiplexed on the MiSeq and raw reads were quality checked using FastQC (Andrews, 2010). Primers were removed from the sequences using cutadapt (Martin, 2011) with a single mismatch allowed (parameters: -e 0.05 --discard-untrimmed). The reads were subsequently processed using the DADA2 package (Callahan et al., 2016) within R (R Team, 2018). Reads were truncated (165 and 160 bp for forward and reverse reads respectively) and filtered with a maximum allowable number of "expected errors" (maxEE) of four (forward reads) and six (reverse reads) with the exception of the Red Sea where sequence quality allowed a more stringent criteria of two and four for the forward and reverse reads respectively. A parametric error matrix was constructed using the first 10^8 bp of the sequences from each sequencing run. Sequences were dereplicated and sequence variants inferred based on this error matrix with singletons being discarded. The remaining paired-end reads were merged with a minimum overlap of 10 bp and a maximum mismatch of 0 bp. At this point, all sequencing runs were combined and chimeric sequences were removed using the removeBimeraDenovo script within DADA2. Pseudogenes were detected and removed using the methods described in Leray and Knowlton (2015). Sequences were translated and aligned against a subset of the MIDORI database (Machida, Leray, Ho, & Knowlton, 2017) using Multiple Alignment of Coding Sequences (MACSE; (Ranwez, Harispe, Delsuc, & Douzert, 2011)). The translations were first made using the invertebrate code and then the vertebrate code. Any sequences with a stop codon or possessing greater than two frame shifts were considered as pseudogenes and removed from further analysis. The number of reads removed at each step is depicted in Supplementary Table 2. The scripts used for processing the reads can be found at: https://github.com/jkpearmanbioinf/DEVOTES-Bioinformatics

Reference sequences of the ASVs were taxonomically classified using the method implemented by the Ribosomal Database Project (rdp) as described in (Wang, Garrity, Tiedje, & Cole, 2007) against a redundant BOLD database (Ratnasingham & Hebert, 2007) supplemented with COI sequences from NCBI. The threshold for taxonomic assignment was 0.51 and ASVs that were assigned at the kingdom level to eukaryotes kept. For comparisons between samples and fractions all samples were subsampled to an even depth of 10000 sequences (meaning that a total ARMS sample comprised 30000 sequences over the three fractions). For bioinformatics analysis, the fractions of the ARMS (sessile; mobile: 106 – 500 μm and 500 – 2000 μm) were combined in two different ways. Firstly, the total ARMS was considered and this comprised all three fractions being merged together. If a fraction did not meet the threshold of subsampling then this whole ARMS was removed from the analysis. A second grouping of fractions separated the sessile from the mobile assemblages. The mobile organisms (106 – 500 μm and 500 – 2000 μm fractions) were merged to comprise a single mobile fraction (combined mobile samples comprised 20000 sequences per ARMS). Similarly, fractions with lower values were not considered further.

Alpha diversity statistics were calculated within *phyloseq* (McMurdie & Holmes, 2013) and tested for significant differences (One way Analysis of Variance (ANOVA) on sqrt transformed data) for the factor region. Tests for normality (shapiro.test) and homogeneity of variance (leveneTest; package car (Fox & Weisberg, 2019)) were undertaken prior to running the ANOVA. Post-hoc tests were undertaken using the Tukey HSD method. The ASV richness was correlated against environmental variables (see below for how they were selected) using Spearman's Rank correlations. To assess the taxonomic composition of the communities in each region, ASVs were combined at the phylum level using *phyloseq* in R. The composition was depicted via donut plots with *ggplot* (Wickham, 2016) and on bathymetric maps produced with *marmap* (Pante & Simon-Bouhet, 2013). PCoA plots were produced using square root transformed community data with distances matrices constructed using both Bray Curtis and Jaccard dissimilarity functions and Euclidean distances for the environmental variables. The PCoA ordination was undertaken in *phyloseq* and plotted in *ggplot*. Statistical differences were tested using PERMANOVA undertaken in the adonis function of the R package vegan (Oksanen et al., 2007) on the square root transformed data (one factor: Region; 6 levels:). Pairwise permutation tests were undertaken with RVAidememoire (Hervé, 2017). Local Contributions to Beta Diversity (LCBD) as described by Legendre & De Caceres (2013), is the degree of uniqueness of the sites

and was calculated in *adespatial* (Dray et al., 2016). Total diversity was calculated on a Hellinger transformed community matrix. Beta diversity was partitioned computing the sums of squares of a sampling unit as a proportion of the total diversity (Legendre & De Caceres, 2013). Richness and LCBD were compared to see if there was a correlation between richness and the contribution of each site to beta diversity.

To disentangle whether ecological niche factors or dispersal processes are more important explaining species composition across seas, we related beta diversity indices with environmental factors and geographic distance among seas, and among sites and replicates. To address this, similarity/distance matrices were created by investigating pair-wise comparisons of all replicates, sites and seas for three categories: biotic similarity, environmental distance, and geographic distance. For the biotic similarity matrix, we calculated pairwise community similarities using the Jaccard dissimilarity index across sites (Legendre & Legendre, 2012) with ASV presence/absence data at each station. Environmental distance was computed pair-wise among replicates, sites and seas based on Euclidean distance using five environmental variables obtained from the Bio-oracle database (Assis et al., 2018; Tyberghein et al., 2012) within the 2000-2014 period and extracted for each biological station: Sea Surface Salinity (SSS), Chlorophyll-a concentration (Chl-a), Sea Surface Temperature (SST), Sea Surface Temperature Range (SST-range), Photosynthetic Available Radiation (PAR) (Supplementary Table 3). Environmental variables selected represent limiting factors, resources, or disturbances (natural or anthropogenic) causally linked to the marine benthic species and its habitat (Reiss et al., 2015); salinity provides information on freshwater sources and it is known to be low in the Black Sea and Baltic Sea compared to other European seas; chlorophyll-a concentration characterize productive seas and anthropogenic eutrophic areas; sea temperature is one of the main influencing factor to marine species; sea temperature range is a proxy to characterize seas with thermal extremes; diffusion attenuation coefficient is a proxy of water transparency which limits primary producers; and PAR is used to measure the spectral range of light that is available in the water column for use by algae for photosynthesis. The selection of those variables was also based on various criteria: availability in a standardized way (i.e. in Bio-oracle), avoiding other variables correlated with those selected, previous knowledge of reliable variables (most of them combine satellite and *in situ* data, hence are better estimated than those that cannot be obtained from satellite data such as nutrients, for instance, that can have important local spatial variability), and knowledge of influence on benthic species. The geographic distance

matrix was calculated as the shortest oceanic path between two sampling sites (km), avoiding land. To do so, we used the marmap (Pante & Simon-Bouhet, 2013) package in R.

Prior to the analysis relating biotic and environmental variables, we tested for collinearity between explanatory variables by calculating variance inflation factors (VIF) with the *ade* package in R (Zuur, Ieno, Walker, Saveliev, & Smith, 2009). We excluded any variable that had a VIF > 5, and then recalculated VIF for the remaining variables. We iterated this process until all variables had a VIF < 5. Mantel and partial Mantel correlations (Legendre & Legendre, 2012) were calculated between ASV similarity and both environmental and geographic distances to analyze the relative contribution of niche and space descriptors to benthos community structure, respectively. Mantel tests were calculated using the *vegan* package. The slope of the line between assemblage similarity and either spatial or environmental distance was calculated with linear regression and used as an indicator for the rate of spatial species turnover among assemblage composition (distance-decay). Steeper negative slopes indicate faster spatial turnover, whereas slope = 0 indicates no turnover over space. We also calculated the halving-distance metric that identifies the distance at which community similarity halves, and provides relevant information regarding the spatial scale of community variation (Soininen, McDonald, & Hillebrand, 2007). Halving-distance for each community (sessile vs mobile) were calculated following Villarino et al., (2018).

Results

Local diversity patterns

Average ASV richness per ARMS varied significantly across the six regions (ANOVA: F value = 47.62; $p < 0.001$, df=5; residual df =30), ranging between 360 and 1615 ASVs in the Black and Red Seas, representing 1.5% and 6.8% of the gamma diversity (23922 ASVs including only those ARMS that had all fractions, 27473 ASVs total), respectively (Supplementary Table 4). Red Sea ASV richness was significantly ($p < 0.001$) higher than in all the other regions except for the Mediterranean Sea ($p = 0.68$). The Baltic and Adriatic Seas presented a similar number of taxa to the Bay of Biscay ($p = 0.34$ and $p = 0.41$ respectively), with the Black Sea having significant differences to the other regions ($p < 0.001$) (Figure 1). There was a significant difference in the Shannon diversity across seas (Kruskal-Wallis chi-squared = 24.44, df = 5, p-value < 0.002). The Baltic and Black Sea had a significantly (all comparisons $p < 0.03$) lower Shannon diversity than the other seas.

No clear latitudinal or longitudinal gradient was apparent from our dataset in terms of ASV richness. However, a negative Spearman's rank relationship with the range of SST (Spearman's Rank rho = -1, p = 0.003) and Chl-*a* (Spearman's Rank rho = -0.886, p = 0.033), and a positive relationship with SSS (Spearman's Rank rho = 0.886, p = 0.033) were observed. For Shannon diversity we observed a negative correlation with the range of SST (Spearman's Rank rho = -0.942, p = 0.016) and Chl-a (Spearman's Rank rho = -0.942, p = 0.016). The two regions with the lowest number of observed ASVs also had the lowest local contribution to beta diversity (Figure 2A). There was a general positive correlation (Linear regression: Total: $R^2 = 0.4513$; $p < 0.001$, Sessile: $R^2 = 0.2852$; $p = < 0.001$, Mobile: $R^2 = 0.4666$; $p < 0.001$) between the number of observed ASVs and LCBD across regions (Figure 2B).

Approximately 17% of the ASVs were shared between the sessile and mobile groups. No ASVs were present in all the regions. Indeed, the number of ASVs present in the Red Sea and in at least one of the other regions was generally low despite the Red Sea having the highest richness (Supplementary Table 5). Taxonomically these shared ASVs were within the groups Sessilia (2 ASVs) and Primates (7 ASVs) as well as unclassified ASVs (Supplementary Table 6). The ASVs belonging to Primates, were attributed to humans and are likely to be contamination and are not considered further. The highest number of shared ASVs was observed between the Western Mediterranean and the Bay of Biscay (covering 12 classes) and the Adriatic Sea (covering 13 classes). The two regions with low salinities (Baltic and Black Sea) also had a relatively high number of shared ASVs with the majority of these being related to the barnacle genus *Amphibalanus* (Supplementary Table 6).

The taxonomic composition of the benthic communities changed across the regional seas. The Western Mediterranean Sea and Bay of Biscay had the highest number of phyla (19) whereas the lowest number was observed in Baltic Sea (16 phyla) (Supplementary Table 4). Unique genera were investigated for each region. In the Red Sea three of the most abundant genera were groups of reef fish (*Chromis, Pseudochromis and Eviota*) (Supplementary Table 7). In the Bay of Biscay, the three most abundant genera are Crustacea, including two Decapoda (*Pisidia* and *Eurynome*), and one Cirripede (*Verruca*). In the Adriatic Sea, the most abundant unique genera belong to Bryozoa (*Schizoporella*) and to Mollusca, including Bivalvia (*Polititapes*) and Gastropoda

(*Hexaplex*) (Supplementary Table 7). Several of the unique genera in the Western Mediterranean belonged to the red algae (class: Florideophyceae) including *Womersleyella, Peyssonnelia, Kallymenia* and *Lithophyllum* (Supplementary Table 7).

Overall, Arthropoda dominated the cryptobiome in all regions investigated (Figure 1 and Supplementary Table 4) although there were statistical differences between the regions (ANOVA: $F = 14.99$; $p < 0.001$) with the Baltic Sea having a higher relative abundance of reads ($p < 0.001$) than the other seas. Differences were observed in the contribution of the other main phyla with Porifera (14%; ANOVA: $F = 14.51$; $p < 0.001$) having a higher abundance in the Red Sea ($p < 0.01$) especially in the sessile assemblage. Annelida (61%) dominated the sessile benthos in the Adriatic Sea, while Bryozoa were predominant in the Black Sea (sessile: 42%; mobile: 5%). Community composition in the Bay of Biscay and Western Mediterranean Sea was highly comparable, with Rhodophyta, Annelida and Arthropoda being the major constituents. No statistical difference was observed for Annelida and Arthropoda ($p = 0.52$ and $p = 0.99$ respectively) while Rhodophyta was slightly higher in the Bay of Biscay ($p < 0.01$). In general, the contribution of unclassified ASVs to the gamma diversity was substantial (range: 25% - 41%) in the ARMS as a whole, whilst it was lower in terms of sequence abundance (range: 15% - 28%), especially in the sessile biota of the Adriatic Sea (4.6%) and Baltic Sea (3.5%).

Cross-regional biodiversity

Multivariate analysis performed using Jaccard (presence/absence) dissimilarity showed significant spatial differences for both sessile (PERMANOVA: $F=4.026$; $p < 0.001$) and mobile (PERMANOVA: $F = 3.5266$; $p < 0.001$) assemblages (Figure 3). Pairwise permutational tests indicated that the communities in all seas were significantly different for both the sessile and mobile assemblages. Our results also show that differences between communities, can be well mirrored in the environmental PCoA analysis, where environmental variables clustered in well-defined groups according to the different environmental characteristics of regional seas across Europe (Figure 3).

Spatial patterns of benthos beta-diversity are slightly more determined by the oceanic distance among pair sites relative to environmental differences as seen in the partial mantel correlations between community similarity, oceanic distance, and environmental distance (Table1). Higher

associations were observed between community assembly and oceanic distance than to environment are also observed for each taxonomic group, with Chordata, Annelida and Echinodermata showing the highest partial correlations (Table 2, Figure 4). Thus, for these latter groups, dispersal limitation is particularly important in driving spatial distribution (Table 2, Figure 4). Community similarities decreases with oceanic distance showing a similar pattern (turnover rate, slope of distance-decay) in both mobile and sessile fractions (Table 3, Figure 5). Community similarity decay against environmental differences is also similar for both groups (slopes = -0.1986 and -0.1936 for the sessile and mobile respectively) (Figure 5). Indeed, the differences in environmental distances correlates well with oceanic distance (r^2=0.55). Mantel tests reveal that community similarities are significantly correlated with all environmental descriptors, especially SST, SST-range and PAR (Figure 6). The Mantel correlation for each of the environmental variable, as well as the sum of all of them, is slightly higher in the mobile group compared to the sessile one. A similar environmental decay response is observed for both the sessile and mobile groups (Figure S1) with temperature (mean and range), having steeper slopes than with other set environmental variables.

Discussion

Our metabarcoding-based pan-regional study provides a framework for the assessment of the ARMS biodiversity associated with rocky-bottoms and their patterns, including those of the often neglected "hidden majority" of small organisms contributing to the cryptic community (Pearman et al., 2018; Ransome et al., 2017). Here, we reveal pan-regional patterns in diversity of benthic organisms with the marginal seas (i.e. Baltic, Black and Red Seas) having a distinct composition, similar to that found using other methods (Cahill et al., 2018; David et al., 2019).

Pan-regional biodiversity patterns

No clear trends in richness or Shannon diversity were observed with latitude or longitude with the highest values being observed in the Red Sea and the lowest ones in the Black Sea. However, there were significant trends with environmental variables. Negative associations were observed between richness and diversity and the range of SST and chlorophyll *a*. Further SSS showed a positive relationship with richness and diversity. Temperature has often been suggested as having an impact on diversity primarily due to the energy hypothesis where increased metabolic rates promote higher speciation (Tittensor et al., 2010). However, environmental stability, as indicated

here by lower SST ranges has also been previously shown to have negative associations with richness (Rombouts et al., 2009; Tittensor et al., 2010). The Red Sea and Western Mediterranean both had high values of ASV richness. ARMS were deployed on coral reef and photophilic algae biocenosis habitats respectively in the Red and Western Mediterranean Seas, respectively and it is likely that these habitats have higher heterogeneity than other temperate hard bottomed communities (Loreau, 2000). The higher structural complexity would provide more ecological niches, attenuating the effects of competition and promoting the co-existence of more species at the scale of the reef (Gratwicke & Speight, 2005).

In general, the contribution of each site to beta diversity (LCBD) was positively correlated with richness. This is especially noticeable for the Red Sea which, for both the sessile and mobile groups, has comparatively high values for both richness and LCBD compared with other regions. This suggests that the Red Sea has a distinct community relative to the other seas considered here, which is backed up with the identification of a limited number of ASVs that are shared between the Red Sea and the other seas and is a similar trend to that observed by Cahill et al. (2018) using artificial substrate units (ASUs). This is likely a reflection of the predominant Indo-Pacific origin of the Red Sea fauna (Sherman, Okemwa, & Ntiba, 2009) with only the artificial Suez Canal, built between 1859 and 1869, connecting the Red Sea with the Mediterranean Sea. In the current study, the order Sessilia (2 ASVs) and 9 ASVs unidentified could be identified that were shared between the Red Sea and the other seas. The first group was in the order Sessilia, the acorn barnacles, which was observed in both the Red Sea and the Adriatic. While it was not possible to identify the ASVs below the order Sessilia, species in this order have been observed in both the Adriatic and the Red Sea (Igić, 2007).

Interestingly, a relatively high number of ASVs were shared between the Baltic and Black Seas and this may account for the low values of LCBD in these regions as they were also the poorest in terms of richness. The majority of these ASVs were taxonomically attributed to *Amphibalanus improvisus.* This species is one of the most successful aquatic invaders and is proposed to spread via shipping routes (Wrange et al., 2016). While *Amphibalanus improvises* has a broad environmental tolerance it performs best at low salinity (Wrange et al., 2014). While it is likely that this species has been transported throughout European waters through shipping routes, the niche preference of this species for low salinity environments could explain why it is

predominantly found in the Baltic and Black Seas in the current dataset. The Black Sea shared considerably fewer ASVs with the Adriatic and Western Mediterranean than it did with the Baltic. The Black Sea is connected to the Mediterranean through the Bosphorus strait. Prior to the Bosphorus connection being formed approximately 7000 years ago, there is evidence of long-term isolation of fauna dating from the Pleistocene (Nikula & Väinölä, 2003; Papadopoulos, Peijnenburg, & Luttikhuizen, 2005). About 7000 years ago, connection to the Mediterranean Sea, through the Bosphorus, led to an increase in the salinity of the Black Sea, allowing for the introduction of species from the Mediterranean (Zaitsev & Mamaev, 1997); however, the relatively lower salinity still present, as well as the colder winter temperatures may act as an ecological barrier that differentiates the communities within the two regions.

The highest number of ASVs were shared between the Western Mediterranean and the Adriatic (116 ASVs) and the Bay of Biscay (115 ASVs). Despite the higher oceanic distance, the Bay of Biscay and Western Mediterranean populations shared similar numbers of ASVs compared with the Western Mediterranean and Adriatic Sea which are geographically closer. This could indicate a physical barrier to dispersal being present, caused by the Sicilian-Tunisian strait that has been implicated in limited gene flow between West and East Mediterranean populations of fish, invertebrates and seagrasses (Arnaud‐Haond et al., 2007; Bahri-Sfar, Lemaire, Hassine, & Bonhomme, 2000; Borsa et al., 1997; Debes, Zachos, & Hanel, 2008). A complementary explanation could come from the environmental similarity between Western Mediterranean Sea and Bay of Biscay, since the Adriatic presents different environmental conditions (Supplementary Table 3) such as higher diffuse attenuation values, chlorophyll *a* concentration and sea temperature range, with hotter summers and colder winters, according to our results and described elsewhere (Borja et al., 2004; Cahill et al., 2018). The fact that only a single ASV (belonging to the genus *Oscarella* (Porifera)) was found in all three regions in this dataset indicated that a combination of the variables described above was limiting the number of species found in these locations that shared a similar latitudinal position.

The Western Mediterranean Sea and Bay of Biscay had the highest number of phyla (19) whereas the lowest number was observed in the Baltic Sea (16 phyla). While this study is unlikely to be a comprehensive assessment of the phyla present in the regions, due to limited within region replication, possible primer bias, as well as incomplete reference molecular databases (Danovaro

et al., 2016), it confirms the rich phylogenetic composition of the Western Mediterranean, Bay of Biscay, and the Red Sea (18 phyla), which have previously been reported as hotspots of diversity (Coll et al., 2010). Overall, arthropods dominated the cryptobiome across all seas. The phylum Arthropoda, and particularly the class Crustacea, is known to have the greatest species richness in the marine system (Narayanaswamy et al., 2013). Changes in the relative contribution of Arthropoda to the overall biodiversity were aligned with the latitudinal extremes, being highest in the Baltic Sea and at a minimum in the Red Sea. Community composition in Bay of Biscay and Western Mediterranean Sea was highly comparable, with Rhodophyta, Annelida and Arthropoda being the major constituents, with these two regions sharing the highest percentage of ASVs with this observation similar to previous non ARMS studies (Borja et al., 2004, 2019; Cahill et al., 2018; Wangensteen & Turon, 2015).

The most abundant unique genera in the Red Sea (*Chromis, Pseudochromis* and *Eviota*) are Indo-Pacific fish taxa which are commonly found on coral reefs within the Red Sea. In the Adriatic Sea, the unique genera were dominated by Bivalvia and Gastropoda with these groups often dominating the macrobenthic community on artificial substrates this region (Cahill et al., 2018, 2018; Spagnolo et al., 2014). Species from the genus *Polititapes*, especially *P. aureus*, can be described as common bivalve in this basin (Milišić, 2007). Further, *Hexaplex* is one of the most abundant and widespread genera of muricid gastropods in the Adriatic, and some species, such as *H. trunculus*, are fished for human consumption and used as fish bait (Benović, 1997). The Bryozoan genera *Schizoporella* are commonly observed in the Adriatic Sea (Hayward & McKinney, 2002) and is tolerant to high salinity and mostly settle on hard bottoms of various origins (rocks, shells, logs, algae, sea plants etc) (Igić, 2007). ARMS were positioned in a habitat where red algae were predominant in the Western Mediterranean so the presence of several unique genera belonging to the class Florideophyceae would be expected with *Verlaquea* (synonym of *Kallymenia*) *lacerata* for example being common in the western parts of the Mediterranean and present in the sampling area, but rare elsewhere in the basin (Rodríguez-Prieto & Vergés, 2001).

Relative contribution of niche and dispersal processes structuring the cryptobiome

Dispersal limitation and niche descriptors determine community beta-diversity patterns in rocky shore benthos to a similar amount, although values for dispersal limitation are slightly higher both

in general and for each one of the five taxonomic groups, in particular if we look at the partial Mantel correlations (Oceanic distance (dispersal): patial r_M=0.25-0.43; Environmental (niche): partial r_M=0.07-0.26). A previous study in a soft-bottom macroinvertebrate community using traditional taxonomy methods revealed dispersal traits to be more determinant on benthic regional beta-diversity patterns (Chust et al., 2016). The use of metabarcoding in the current study enabled a broader range of taxa to be incorporated into the analysis including those with smaller body sizes and different life histories. While dispersal limitation was observed to be a slightly stronger determinant, a high shared covariation between environmental distance and spatial distance is observed, indicating that both descriptors might contribute to structuring the community, making it difficult to estimate their exact relative importance.

That no differences in the species turnover rate (slope of distance decay) and dispersal scale (halving distance) between biological groups found is also worth mentioning (Table 3). That sessile organisms often have a larval dispersal phase (Shanks, 2009) may explain the similarity in these dispersal scales. This might indicate that both sessile and mobile groups have similar dispersal constraints from local to regional scales. Additionally, the portion of ASVs shared between sessile and mobile groups (17%) could partially explain this pattern. In contrast, the correlation between species similarity and environmental variables is slightly higher in the mobile group compared to the sessile one, which is interpreted as the sessile group having a wider niche breath.

While investigating species was beyond the ability of the current study analysis was conducted for the five most ASV rich phyla. For all phyla the contribution of oceanic distance and environmental distance to the distribution was similar with oceanic distance in general being slightly higher. In the case of Chordata, environmental contribution is not significant, indicating that dispersal limitation might strongly constrain their distribution. However, we note the limitations of investigating the response at a phylum level with a variety of responses likely to occur within each phylum depending on the species.

There are some limitations to detecting barriers to dispersal based on relating community similarity with geographical distances. Dispersal of marine organisms through the seascape is mainly determined by the ocean currents and the dispersal capacity of organisms (e.g. Alberto et

al., 2011; Cowen, Gawarkiewicz, Pineda, Thorrold, & Werner, 2007). Therefore, a better estimate of oceanic connectivity than oceanic distance (as calculated by geographic distance circumventing land) would be the surface ocean transit times from site to site, derived for instance from Lagrangian particle simulations (Jönsson & Watson, 2016; Villarino et al., 2018). However, this approach needs global current models to cover all the assessed seas, and these models are unavailable at the spatial resolution needed in our case, for instance connections through the Strait of Gibraltar, Bosphorus and Suez Canal. Others constraints are as follows: 1) the current ARMS sampling scheme is not comprehensive enough for abrupt transitions to be detected; 2) ARMS sampling does not consider temporal variation; hence, we do not have information on the colonization rates nor the dynamics of dispersal rates; and 3) tracking the trajectory and fate of propagules (e.g. Selkoe & Toonen, 2011; Shanks, 2009) are methods that enable direct measures to detect dispersal barriers; however, deploying these methodologies to cover the overall biogeographic area we are considering is not feasible.

Limitations of ARMS and molecular methods to assess the hard-bottom benthic environment.

This study represents one of the largest cross-seas metabarcoding surveys of the cryptobiome conducted to date. As a consequence, we were able to advance the understanding of the drivers of benthic cryptobiome community structure. However, we do recognize that there are potential limitations to the methodologies used. Firstly, ARMS are artificial substrates and thus may have a different colonization pattern to that of the natural substrates. Limited comparisons have been undertaken but Plaisance et al., (2011) for invertebrates and Pennesi and Danovaro (2017) for the microphytobenthos have shown similarities with natural communities. Further comparisons between ARMS and natural substrates in a variety of habitats would be required, especially if ARMS were to be incorporated into long term monitoring systems, to assess truly how representative the ARMS community is compared to the natural habitat. We also recognize that despite using up to date bioinformatic pipelines, the molecular approach used have some clear limitations (Zinger et al., 2019) and, therefore, the results should be considered carefully. Firstly, samples were rarefied at 10000 reads and this may not reveal their full diversity especially in species rich areas. To get a better estimation of diversity further replicates could be taken (Lanzén, Lekang, Jonassen, Thompson, & Troedsson, 2017). In the current study, sequencing was undertaken at several facilities which could possibly lead to batch effects due to differences in

library preparation. While error matrices were constructed for each run individually the possibility of biases could still exist. While it was not undertaken in this study, the construction of mock communities could highlight differences in the sequencing efficiencies among facilities (Deiner et al., 2017). Negative controls were undertaken in this study and were investigated by gel electrophoresis and shown to be negative. However, there is a possibility that contamination could be below the level of detection on a gel and thus, there is a possibility that a low level of diversity within the samples could be due to contamination (Deiner et al., 2017). Challenges also persist in estimations of relative abundance due to a mixture of field and laboratory factors (Kelly, 2016). This can be due to differences in the production of eDNA (Klymus, Richter, Chapman, & Paukert, 2015), the production of co-purifying contaminants (e.g. mucopolysaccharides) by taxa (Pereira, Chaves, Bastos, Leitão, & Guedes-Pinto, 2011), and also through primer biases that skew these estimates (Elbrecht & Leese, 2015). Indeed, this could explain the relative lack of molluscs in the current dataset. This low abundance was also found by Cahill et al., (2018) in metabarcoding results which contrasted with morphological results undertaken at the same sites as this study but using ASUs instead. Furthermore, a large proportion of the ASVs could not be given reliably phylogenetic assignments even at phylum level. Improvements in taxonomic assignments would enhance the analysis of the biological groups and could allow the analysis to be undertaken at lower phylogenetic levels (e.g. class or family). A multi-gene approach to assessing the communities would likely mitigate primer bias and to a certain extent taxonomic assignment issues and reveal a more complete view of the communities. However, it has previously been shown that while complementary patterns were observed for taxonomic compositions, the spatial patterns observed for the two genes (18S rRNA and COI) were similar (Pearman et al., 2018). Although clear differentiations amongst regions were observed and the study covers large spatial scales; we believe further sampling within each region would give a better estimation of the effects of ocean distance and environmental distance across multiple scales and would also improve estimations of connectivity between regions. Lastly, communities are known to change over time and the current study only investigated the cumulative community present on the ARMS over an annual period. Looking at the community changes over multiple temporal scales (e.g. seasonal to interannual) would give a better idea of how communities change in and amongst different regions and can also provide some baseline data to the taxa regularly present in the community.

Conclusions

By using a standardized sampling and analysis protocol, we show that oceanic distance among sampling sites are slightly more important than niche descriptors in driving ASV distribution. However, high-shared covariation between environmental distance and oceanic distance is observed, indicating that both descriptors might contribute to structuring the rocky bottom mobile and sessile communities, and making it difficult to estimate their relative importance. While understanding spatial patterns in biodiversity is essential to fully discern anthropogenic impacts on ecosystems, temporal patterns via long-term datasets, are also critical (Magurran et al., 2010). However, to avoid bias in assessments over time, which often involve changing monitoring equipment, quantifications and characterizations of biodiversity should be standardized. Here, we have presented a study across multiple institutions spanning different regional seas using standardized ARMS and a metabarcoding approach. We believe that this framework has the potential to improve the monitoring of rocky areas at global scale. For instance, long term insights are crucial to understanding ecological processes (Seddon et al., 2014) and the retrieval and replacement of ARMS in the exact same spot has the potential to create comparable long-term datasets. With ARMS units deployed temporally at local, regional and global scales investigations into the changes in the presence and absence of species across time may be feasible, which has the potential to show shifts in the species with global pressures such as climate change. This approach will provide managers and other stakeholders valuable information about how the cryptobiome – comprised of small and highly diverse, yet often neglected organisms – respond to a variety of natural and anthropogenic impacts. This information can only benefit policy making, as it incorporates responses of organisms that have different life histories and functions than the macro-organisms (e.g. algae, macroinvertebrates and fish) that are currently relied on.

Acknowledgements

This manuscript is a result of the DEVOTES (DEVelopment Of Innovative Tools for understanding marine biodiversity and assessing good Environmental Status) project, funded by the European Union under the 7th Framework Programme, "The Ocean of Tomorrow" Theme (grant agreement no. 308392), www.devotes-project.eu. S Carvalho and JK Pearman were funded through the Saudi Aramco-KAUST Center for Marine Environmental Observations (SAKMEO). MC Uyarra was partially funded through the Spanish Programme for Talen and Employability in

R + D + I “Torres Quevedo”. We would also like to thank the editor Prof. Creer and three anonymous reviewers for the time and effort they dedicated to reviewing the manuscript.

References

Alberto, F., Raimondi, P. T., Reed, D. C., Watson, J. R., Siegel, D. A., Mitarai, S., … Serrão, E. A. (2011). Isolation by oceanographic distance explains genetic structure for Macrocystis pyrifera in the Santa Barbara Channel. *Molecular Ecology*, *20*(12), 2543–2554. doi: 10.1111/j.1365-294X.2011.05117.x

Al-Rshaidat, M. M., Snider, A., Rosebraugh, S., Devine, A. M., Devine, T. D., Plaisance, L., … Leray, M. (2016). Deep COI sequencing of standardized benthic samples unveils overlooked diversity of Jordanian coral reefs in the northern Red Sea. *Genome*, *59*(9), 724–737. doi: 10.1139/gen-2015-0208

Andersen, J. H., Halpern, B. S., Korpinen, S., Murray, C., & Reker, J. (2015). Baltic Sea biodiversity status vs. cumulative human pressures. *Estuarine, Coastal and Shelf Science*, *161*, 88–92.

Andrews, S. (2010). *FastQC: a quality control tool for high throughput sequence data*.

Arnaud-Haond, S., Migliaccio, M., Diaz-Almela, E., Teixeira, S., Vliet, M. S. V. D., Alberto, F., … Serrão, E. A. (2007). Vicariance patterns in the Mediterranean Sea: east–west cleavage and low dispersal in the endemic seagrass Posidonia oceanica. *Journal of Biogeography*, *34*(6), 963–976. doi: 10.1111/j.1365-2699.2006.01671.x

Assis, J., Tyberghein, L., Bosch, S., Verbruggen, H., Serrão, E. A., & Clerck, O. D. (2018). Bio-ORACLE v2.0: Extending marine data layers for bioclimatic modelling. *Global Ecology and Biogeography*, *27*(3), 277–284. doi: 10.1111/geb.12693

Bahri-Sfar, L., Lemaire, C., Hassine, O. K. B., & Bonhomme, F. (2000). Fragmentation of sea bass populations in the western and eastern Mediterranean as revealed by microsatellite polymorphism. *Proceedings of the Royal Society of London B: Biological Sciences*, *267*(1446), 929–935.

Barbier, E. B., Hacker, S. D., Koch, E. W., Stier, A. C., & Silliman, B. R. (2011). *Estuarine and coastal ecosystems and their services*.

Benović, A. (1997). The history, present condition, and future of the molluscan fisheries of Croatia. In *The history, present condition, and future of the molluscan fisheries of North and Central America and Europe* (pp. 217–226). Seattle, Washington, USA: U.S. Department of Commerce.

Bestová, H., Munoz, F., Svoboda, P., Škaloud, P., & Violle, C. (2018). Ecological and biogeographical drivers of freshwater green algae biodiversity: from local communities to large-scale species pools of desmids. *Oecologia*, *186*(4), 1017–1030. doi: 10.1007/s00442-018-4074-x

Borja, A., Aguirrezabalaga, F., Martínez, J., Sola, J. C., García-Arberas, L., & Gorostiaga, J. M. (2004). Benthic communities, biogeography and resources management. *Oceanography and Marine Environment of the Basque Country*, *70*(18), 455–492.

Borja, A., Amouroux, D., Anschutz, P., Gómez-Gesteira, M., Uyarra, M. C., & Valdés, L. (2019). The Bay of Biscay. In *World Seas: an Environmental Evaluation* (pp. 113–152). London: Academic Press, Elsevier.

Borsa, P., Naciri, M., Bahri, L., Chikhi, L., Garcia de Leon, F. J., Kotoulas, G., & Bonhomme, F. (1997). Infraspecific zoogeography of the Mediterranean: population genetic analysis on sixteen Atlanto-Mediterranean species (fishes and invertebrates). *Vie et Milieu (France)*.

Burgess, S. C., Baskett, M. L., Grosberg, R. K., Morgan, S. G., & Strathmann, R. R. (2016). When is dispersal for dispersal? Unifying marine and terrestrial perspectives. *Biological Reviews*, *91*(3), 867–882.

Cahill, A. E., Pearman, J. K., Borja, A., Carugati, L., Carvalho, S., Danovaro, R., … Olenin, S. (2018). A comparative analysis of metabarcoding and morphology-based identification of benthic communities across different regional seas. *Ecology and Evolution*, *8*(17), 8908–8920.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*(7), 581.

Carvalho, S., Aylagas, E., Villalobos, R., Kattan, Y., Berumen, M. L., & Pearman, J. K. (2019). Beyond the visual: using metabarcoding to characterize the hidden reef cryptobiome.

Proceedings of the Royal Society B: Biological Sciences, *286*(1896), 20182697. doi: 10.1098/rspb.2018.2697

Chave, J. (2004). Neutral theory and community ecology. *Ecology Letters*, *7*(3), 241–253. doi: 10.1111/j.1461-0248.2003.00566.x

Chust, G., Villarino, E., Chenuil, A., Irigoien, X., Bizsel, N., Bode, A., … Borja, A. (2016). Dispersal similarly shapes both population genetics and community patterns in the marine realm. *Sci Rep*, *6*, 28730. doi: 10.1038/srep28730

Coen, L. D. (1988). Herbivory by Caribbean majid crabs: feeding ecology and plant susceptibility. *Journal of Experimental Marine Biology and Ecology*, *122*(3), 257–276.

Coll, M., Piroddi, C., Steenbeek, J., Kaschner, K., Lasram, F. B. R., Aguzzi, J., … Voultsiadou, E. (2010). The Biodiversity of the Mediterranean Sea: Estimates, Patterns, and Threats. *PLOS ONE*, *5*(8), e11842. doi: 10.1371/journal.pone.0011842

Cordes, E. E., Jones, D. O. B., Schlacher, T. A., Amon, D. J., Bernardino, A. F., Brooke, S., … Witte, U. (2016). Environmental Impacts of the Deep-Water Oil and Gas Industry: A Review to Guide Management Strategies. *Frontiers in Environmental Science*, *4*. doi: 10.3389/fenvs.2016.00058

Costanza, R., d'Arge, R., De Groot, R., Farber, S., Grasso, M., Hannon, B., … Paruelo, J. (1997). The value of the world's ecosystem services and natural capital. *Nature*, *387*(6630), 253.

Costello, M. J., & Wilson, S. P. (2011). Predicting the number of known and unknown species in European seas using rates of description. *Global Ecology and Biogeography*, *20*(2), 319–330. doi: 10.1111/j.1466-8238.2010.00603.x

Cowen, R., Gawarkiewicz, G., Pineda, J., Thorrold, S., & Werner, F. (2007). Population Connectivity in Marine Systems: An Overview. *Oceanography*, *20*(3), 14–21. doi: 10.5670/oceanog.2007.26

Danovaro, R., Carugati, L., Marco, B., Cahill, A. E., Carvalho, S., Chenuil, A., … Borja, A. (2016). Implementing and Innovating Marine Monitoring Approaches for Assessing Marine Environmental Status. *Frontiers in Marine Science*, *3*, 213.

David, R., Uyarra, M. C., Carvalho, S., Anlauf, H., Borja, A., Cahill, A. E., … Chenuil, A. (2019). Lessons from photo analyses of Autonomous Reef Monitoring Structures as tools to

detect (bio-)geographical, spatial, and environmental effects. *Marine Pollution Bulletin*, *141*, 420–429. doi: 10.1016/j.marpolbul.2019.02.066

Debes, P. V., Zachos, F. E., & Hanel, R. (2008). Mitochondrial phylogeography of the European sprat (Sprattus sprattus L., Clupeidae) reveals isolated climatically vulnerable populations in the Mediterranean Sea and range expansion in the northeast Atlantic. *Molecular Ecology*, *17*(17), 3873–3888. doi: 10.1111/j.1365-294X.2008.03872.x

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., … Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, *26*(21), 5872–5895. doi: 10.1111/mec.14350

Dray, S., Blanchet, G., Borcard, D., Guenard, G., Jombart, T., Larocque, G., … Wagner, H. (2016). *adespatial: Multivariate Multiscale Spatial Analysis. R package version 0.0-4*.

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*(7), e0130324. doi: 10.1371/journal.pone.0130324

Elliott, M. (2014). Integrated marine science and management: wading through the morass. *Marine Pollution Bulletin*, *1*(86), 1–4.

Enochs, I. C., Toth, L. T., Brandtneris, V. W., Afflerbach, J. C., & Manzello, D. P. (2011). Environmental determinants of motile cryptofauna on an eastern Pacific coral reef. *Marine Ecology Progress Series*, *438*, 105–118.

Finlay, B. J. (2002). Global dispersal of free-living microbial eukaryote species. *Science*, *296*(5570), 1061–1063.

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (3rd ed.). Thousand Oaks, CA: Sage. Retrieved from https://us.sagepub.com/en-us/nam/an-r-companion-to-applied-regression/book246125

Geller, J., Meyer, C., Parker, M., & Hawk, H. (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*(5), 851–861.

Gratwicke, B., & Speight, M. R. (2005). The relationship between fish species richness, abundance and habitat complexity in a range of shallow tropical marine habitats. *Journal of Fish Biology*, *66*(3), 650–667. doi: 10.1111/j.0022-1112.2005.00629.x

Halpern, B. S., Walbridge, S., Selkoe, K. A., Kappel, C. V., Micheli, F., D'agrosa, C., … Fox, H. E. (2008). A global map of human impact on marine ecosystems. *Science*, *319*(5865), 948–952.

Hayward, P. J., & McKinney, F. K. (2002). Northern Adriatic Bryozoa from the vicinity of Rovinj, Croatia. *Bulletin of the American Museum of Natural History*, *2002*(270), 1–139. doi: 10.1206/0003-0090(2002)270<0001:NABFTV>2.0.CO;2

Hazeri, G., Rahayu, D. L., Subhan, B., Sembiring, A., Anggoro, A. W., Ghozali, A. T., & Madduppa, H. H. (2019). Latitudinal species diversity and density of cryptic crustacean (Brachyura and Anomura) in micro-habitat Autonomous Reef Monitoring Structures across Kepulauan Seribu, Indonesia. *Biodiversitas Journal of Biological Diversity*, *20*(5). doi: 10.13057/biodiv/d200540

Hervé, M. (2017). RVAideMemoire: Diverse Basic Statistical and Graphical Functions (Version 0.9-66). Retrieved from https://CRAN.R-project.org/package=RVAideMemoire

Hubbell, S. P. (1997). A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs. *Coral Reefs*, *16*(1), S9–S21. doi: 10.1007/s003380050237

Hubbell, S. P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press.

Hutchinson, G. E. (1957). Concluding Remarks, Cold Spring Harbor Symposium. *Quantitative Biology*, *22*, 415–427.

Igić, L. (2007). Cirripedia of Adriatic. *Studia Marina*, *24*(2), 1–168.

Jönsson, B. F., & Watson, J. R. (2016). The timescales of global surface-ocean connectivity. *Nature Communications*, *7*, 11239. doi: 10.1038/ncomms11239

Kelly, R. P. (2016). Making environmental DNA count. *Molecular Ecology Resources*, *16*(1), 10–12. doi: 10.1111/1755-0998.12455

Kinlan, B. P., & Gaines, S. D. (2003). Propagule Dispersal in Marine and Terrestrial Environments: A Community Perspective. *Ecology*, *84*(8), 2007–2020. doi: 10.1890/01-0622

Klymus, K. E., Richter, C. A., Chapman, D. C., & Paukert, C. (2015). Quantification of eDNA shedding rates from invasive bighead carp Hypophthalmichthys nobilis and silver carp Hypophthalmichthys molitrix. *Biological Conservation*, *183*, 77–84. doi: 10.1016/j.biocon.2014.11.020

Korpinen, S., & Andersen, J. H. (2016). A global review of cumulative pressure and impact assessments in marine environments. *Frontiers in Marine Science*, *3*, 153.

Lanzén, A., Lekang, K., Jonassen, I., Thompson, E. M., & Troedsson, C. (2017). DNA extraction replicates improve diversity and compositional dissimilarity in metabarcoding of eukaryotes in marine sediments. *PLoS ONE*, *12*(6). doi: 10.1371/journal.pone.0179443

Legendre, P., & De Caceres, M. (2013). Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecology Letters*, *16*(8), 951–963. doi: 10.1111/ele.12141

Legendre, P., & Legendre, L. (2012). Chapter 1 - Complex ecological data sets. In P. Legendre & L. Legendre (Eds.), *Developments in Environmental Modelling* (pp. 1–57). Elsevier. doi: 10.1016/B978-0-444-53868-0.50001-0

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 United States of America*, *112*(7), 2076–2081. doi: 10.1073/pnas.1424997112

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., … Machida, R. J. (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. doi: 10.1186/1742-9994-10-34

Loreau, M. (2000). Biodiversity and ecosystem functioning: recent theoretical advances. *Oikos*, *91*(1), 3–17. doi: 10.1034/j.1600-0706.2000.910101.x

Lotze, H. K., Guest, H., O'Leary, J., Tuda, A., & Wallace, D. (2018). Public perceptions of marine threats and protection from around the world. *Ocean & Coastal Management*, *152*, 14–22.

Machida, R. J., Leray, M., Ho, S.-L., & Knowlton, N. (2017). Metazoan mitochondrial gene sequence reference datasets for taxonomic assignment of environmental samples. *Scientific Data*, *4*(1), 1–7. doi: 10.1038/sdata.2017.27

Magurran, A. E., Baillie, S. R., Buckland, S. T., Dick, J. M., Elston, D. A., Scott, E. M., … Watt, A. D. (2010). Long-term datasets in biodiversity research and monitoring: assessing change in ecological communities through time. *Trends in Ecology & Evolution*, *25*(10), 574–582.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*(1), 10–12. doi: 10.14806/ej.17.1.200

McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLOS ONE*, *8*(4), e61217. doi: 10.1371/journal.pone.0061217

Micheli, F., Halpern, B. S., Walbridge, S., Ciriaco, S., Ferretti, F., Fraschetti, S., … Rosenberg, A. A. (2013). Cumulative human impacts on Mediterranean and Black Sea marine ecosystems: assessing current pressures and opportunities. *PloS One*, *8*(12), e79889.

Milišić, N. (2007). *Glavonošci, puževi i školjke Jadrana*. Split: Marjan Tisak.

Mouquet, N., & Loreau, M. (2003). Community patterns in source-sink metacommunities. *The American Naturalist*, *162*(5), 544–557.

Narayanaswamy, B. E., Coll, M., Danovaro, R., Davidson, K., Ojaveer, H., & Renaud, P. E. (2013). Synthesis of Knowledge on Marine Biodiversity in European Seas: From Census to Sustainable Management. *PLOS ONE*, *8*(3), e58909. doi: 10.1371/journal.pone.0058909

Nikula, R., & Väinölä, R. (2003). Phylogeography of Cerastoderma glaucum (Bivalvia: Cardiidae) across Europe: a major break in the Eastern Mediterranean. *Marine Biology*, *143*(2), 339–350.

Normand, S., Ricklefs, R. E., Skov, F., Bladt, J., Tackenberg, O., & Svenning, J. C. (2011). Postglacial migration supplements climate in determining plant species ranges in

Europe. *Proceedings. Biological Sciences*, *278*(1725), 3644–3653. doi: 10.1098/rspb.2010.2769

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M. H. H., Oksanen, M. J., & Suggests, M. (2007). The vegan package. *Community Ecology Package*, *10*, 631–637.

Pante, E., & Simon-Bouhet, B. (2013). marmap: a package for importing, plotting and analyzing bathymetric and topographic data in R. *PloS One*, *8*(9), e73051.

Papadopoulos, L. N., Peijnenburg, K., & Luttikhuizen, P. C. (2005). Phylogeography of the calanoid copepods Calanus helgolandicus and C. euxinus suggests Pleistocene divergences between Atlantic, Mediterranean, and Black Sea populations. *Marine Biology*, *147*(6), 1353–1365.

Pearman, J. K., 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. doi: 10.1016/j.marenvres.2016.04.011

Pearman, J. K., Aylagas, E., Voolstra, C. R., Anlauf, H., Villalobos, R., & Carvalho, S. (2019). Disentangling the complex microbial community of coral reefs using standardized Autonomous Reef Monitoring Structures (ARMS). *Molecular Ecology*, *28*(15), 3496–3507. doi: 10.1111/mec.15167

Pearman, J. K., Leray, M., Villalobos, R., Machida, R. J., Berumen, M. L., Knowlton, N., & Carvalho, S. (2018). Cross-shelf investigation of coral reef cryptic benthic organisms reveals diversity patterns of the hidden majority. *Scientific Reports*, *8*, 8090. doi: https://doi.org/10.1038/s41598-018-26332-5

Pennesi, C., & Danovaro, R. (2017). Assessing marine environmental status through microphytobenthos assemblages colonizing the Autonomous Reef Monitoring Structures (ARMS) and their potential in coastal marine restoration. *Marine Pollution Bulletin*, *125*(1–2), 56–65. doi: 10.1016/j.marpolbul.2017.08.001

Pereira, J. C., Chaves, R., Bastos, E., Leitão, A., & Guedes-Pinto, H. (2011). An Efficient Method for Genomic DNA Extraction from Different Molluscs Species. *International Journal of Molecular Sciences*, *12*(11), 8086–8095. doi: 10.3390/ijms12118086

Plaisance, L., Caley, M. J., Brainard, R. E., & Knowlton, N. (2011). The diversity of coral reefs: What are we missing? *PloS One*, *6*(10), e25026. doi: 10.1371/journal.pone.0025026

Pocheville, A. (2015). The Ecological Niche: History and Recent Controversies. In T. Heams, P. Huneman, G. Lecointre, & M. Silberstein (Eds.), *Handbook of Evolutionary Thinking in the Sciences* (pp. 547–586). Dordrecht: Springer Netherlands. doi: 10.1007/978-94-017-9014-7_26

Ransome, E., Geller, J. B., Timmers, M., Leray, M., Mahardini, A., Sembiring, A., … Meyer, C. P. (2017). The importance of standardization for biodiversity comparisons: A case study using autonomous reef monitoring structures (ARMS) and metabarcoding to measure cryptic diversity on Mo'orea coral reefs, French Polynesia. *PloS One*, *12*(4), e0175066.

Ranwez, V., Harispe, S., Delsuc, F., & Douzert, E. J. P. (2011). MACSE: Multiple Alignment of Coding SEquences Accounting for Frameshifts and Stop Codons. *PLOS ONE*. doi: https://doi.org/10.1371/journal.pone.0022594

Ratnasingham, S., & Hebert, P. D. N. (2007). BOLD: The Barcode of Life Data System (www.barcodinglife.org). *Molecular Ecology Notes*, *7*, 355–364. doi: 10.1111/j.1471-8286.2006.01678.x

Reaka, M. L. (1987). Adult-juvenile interactions in benthic reef crustaceans. *Bulletin of Marine Science*, *41*(2), 108–134.

Reaka-Kudla, M. I. (1997). The Global Biodiversity of Coral Reefs: A comparison with rain forests. In *Biodiversity II: Understanding and Protecting our Biological Resources* (pp. 83–108).

Reiss, H., Birchenough, S., Borja, A., Buhl-Mortensen, L., Craeymeersch, J., Dannheim, J., … Degraer, S. (2015). Benthos distribution modelling and its relevance for marine ecosystem management. *ICES Journal of Marine Science*, *72*(2), 297–315. doi: 10.1093/icesjms/fsu107

Richter, C., Wunsch, M., Rasheed, M., Kötter, I., & Badran, M. I. (2001). Endoscopic exploration of Red Sea coral reefs reveals dense populations of cavity-dwelling sponges. *Nature*, *413*, 725–730.

Rodríguez-Prieto, C., & Vergés, A. (2001). Geographical Distribution, habitat and reproductive phenology of the genus Kallymenia (Gigartinales, Rhodophyta) from Catalonia, Spain. *Botanica Marina*, *44*(5), 479–492. doi: 10.1515/BOT.2001.054

Rombouts, I., Beaugrand, G., Ibanez, F., Gasparini, S., Chiba, S., & Legendre, L. (2009). Global latitudinal variations in marine copepod diversity and environmental factors.

Proceedings of the Royal Society B-Biological Sciences, *276*(1670), 3053–3062. doi: 10.1098/rspb.2009.0742

Rothans, T. C., & Miller, A. C. (1991). A link between biologically imported particulate organic nutrients and the detritus food web in reef communities. *Marine Biology*, *110*(1), 145–150.

Sammarco, P. W., Kolian, S. R., Warby, R. A. F., Bouldin, J. L., Subra, W. A., & Porter, S. A. (2013). Distribution and concentrations of petroleum hydrocarbons associated with the BP/Deepwater Horizon Oil Spill, Gulf of Mexico. *Marine Pollution Bulletin*, *73*(1), 129–143. doi: 10.1016/j.marpolbul.2013.05.029

Scheffers, S. R., de Goeij, J., van Duyl, F. C., & Bak, R. P. M. (2003). The cave-profilier: a simple tool to describe the 3-D structure of inaccessible coral reef cavities. *Coral Reefs*, *22*, 49–53. doi: 10.1007/s00338-003-0285-6

Seddon, A. W. R., Mackay, A. W., Baker, A. G., Birks, H. J. B., Breman, E., Buck, C. E., … Witkowski, A. (2014). Looking forward through the past: identification of 50 priority research questions in palaeoecology. *Journal of Ecology*, *102*(1), 256–267.

Selkoe, K. A., & Toonen, R. J. (2011). Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. *Marine Ecology Progress Series*, *436*, 291–305. doi: 10.3354/meps09238

Shanks, A. L. (2009). Pelagic larval duration and dispersal distance revisited. *The Biological Bulletin*, *216*(3), 373–385. doi: 10.1086/BBLv216n3p373

Sherman, K., Okemwa, E. N., & Ntiba, M. J. (2009). *Large Marine Ecosystems of the Indian Ocean: Assessment, Sustainability and Management*. John Wiley & Sons.

Soininen, J., McDonald, R., & Hillebrand, H. (2007). The distance decay of similarity in ecological communities. *Ecography*, *30*(1), 3–12.

Spagnolo, A., Cuicchi, C., Punzo, E., Santelli, A., Scarcella, G., & Fabi, G. (2014). Patterns of colonization and succession of benthic assemblages in two artificial substrates. *Journal of Sea Research*, *88*, 78–86. doi: 10.1016/j.seares.2014.01.007

Team, R. C. (2018). *R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014*.

Tittensor, D. P., Mora, C., Jetz, W., Lotze, H. K., Ricard, D., Berghe, E. V., & Worm, B. (2010). Global patterns and predictors of marine biodiversity across taxa. *Nature*, *466*(7310), 1098–1101. doi: 10.1038/nature09329

Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F., & De Clerck, O. (2012). Bio-ORACLE: a global environmental dataset for marine species distribution modelling: Bio-ORACLE marine environmental data rasters. *Global Ecology and Biogeography*, *21*(2), 272–281. doi: 10.1111/j.1466-8238.2011.00656.x

Villarino, E., Watson, J. R., Jönsson, B., Gasol, J. M., Salazar, G., Acinas, S. G., … Chust, G. (2018). Large-scale ocean connectivity and planktonic body size. *Nature Communications*, *9*(1), 142. doi: 10.1038/s41467-017-02535-8

Wang, Q., Garrity, G. M., Tiedje, J. M., & Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, *73*(16), 5261–5267.

Wangensteen, O. S., & Turon, X. (2015). Metabarcoding techniques for assessing biodiversity of marine animal forests. *Marine Animal Forests: The Ecology of Benthic Biodiversity Hotspots*, 1–29.

Weersing, K., & Toonen, R. J. (2009). Population genetics, larval dispersal, and connectivity in marine systems. *Marine Ecology Progress Series*, *393*, 1–12. doi: 10.3354/meps08287

Wickham, H. (2016). *ggplot2: elegant graphics for data analysis*. Springer.

Wilson, E. O., & MacArthur, R. H. (1967). The theory of island biogeography. *Princeton, NJ*.

Wrange, A.-L., André, C., Lundh, T., Lind, U., Blomberg, A., Jonsson, P. J., & Havenhand, J. N. (2014). Importance of plasticity and local adaptation for coping with changing salinity in coastal areas: a test case with barnacles in the Baltic Sea. *BMC Evolutionary Biology*, *14*(1), 156. doi: 10.1186/1471-2148-14-156

Wrange, A.-L., Charrier, G., Thonig, A., Alm Rosenblad, M., Blomberg, A., Havenhand, J. N., … André, C. (2016). The Story of a Hitchhiker: Population Genetic Patterns in the Invasive Barnacle Balanus(Amphibalanus) improvisus Darwin 1854. *PLoS ONE*, *11*(1). doi: 10.1371/journal.pone.0147082

Zaitsev, Y., & Mamaev, V. (1997). Marine biological diversity in the Black Sea. *A Study of Change and Decline United Nations Publications New York*, *208*.

Zimmerman, T. L., & Martin, J. W. (2004). Artificial reef matrix structures (ARMS): an inexpensive and effective method for collecting coral reef-associated invertebrates. *Gulf and Caribbean Research*, *16*(1), 59–64.

Zinger, L., Bonin, A., Alsos, I. G., Bálint, M., Bik, H., Boyer, F., … Taberlet, P. (2019). DNA metabarcoding—Need for robust experimental designs to draw sound ecological conclusions. *Molecular Ecology*, *28*(8), 1857–1862. doi: 10.1111/mec.15060

Zuur, A., Ieno, E. N., Walker, N., Saveliev, A. A., & Smith, G. M. (2009). *Mixed Effects Models and Extensions in Ecology with R*. New York: Springer-Verlag. doi: 10.1007/978-0-387-87458-6

Data Availability

Sequences used for this study are deposited in the NCBI short read archive under accession PRJNA557002 (NCBI 2019). The scripts used for processing the reads can be found at: https://github.com/jkpearmanbioinf/DEVOTES-Bioinformatics. Adapter information can be found in the supplementary information

Author Contributions

JKP, SC, AB, NR-E, LC, RD, ACh, AC, SM organized and led the deployment and retrieval in their respective regions. JKP, AC, LC, IM undertook the lab analysis while JKP, SC, GC, EV, JW undertook the data analysis. JKP, SC, GC, EA wrote the manuscript in discussion with Ach, AC, LC and AB. All authors contributed to various drafts of the manuscript.

Figure Legends

Figure 1: *Top) Map indicating the regions investigated with the insert showing the pan-regional composition (sequence abundance) of ARMS for both the sessile and mobile components. Middle) The regional composition for the sessile (inner ring) and mobile (outer rings) assemblages is*

depicted with an individual donut and Bottom) alpha diversity (# ASVs – Amplicon Sequence Variants-) per ARMS (line across the boxplot depicts the median with the diamond highlighting the mean; the lower hinge represents the 25% quantile with the upper hinge at 75%; the whiskers indicate the 95% confidence interval). Note, "other" in the taxonomy list defines various phyla with low relative abundances.

Figure 2: *A) Local contribution to beta diversity (LCBD) of each site. The dotted grey line represents the value where all sites would contribute equally. B) Local contribution to beta diversity against the observed number of Amplicon Sequence Variants (ASVs). Linear regression for 2B: Total: $R^2 = 0.4513$; $p < 0.001$, Sessile: $R^2 = 0.2852$; $p = < 0.001$, Mobile: $R^2 = 0.4666$; $p < 0.001$. Light grey shading represents the standard error of the linear regression. Colors for the points represent the different regions.*

Figure 3: *Principal Coordinate Analysis plot depicting the variability in the cryptobiome composition based on Jaccard similarity matrices of the COI gene. Analysis was undertaken on the full ARMS unit as well as the sessile and the mobile groups (combination of 106-500 µm and 500-2000 µm fractions). Points were colored according to region. Principal Coordinate Analysis (PCoA) showing the variability of environmental variables among regional seas based on Euclidean distances. Analysis was undertaken on the full ARMS unit (combination of sessile 106-500 µm and 500-2000 µm fractions). Points were colored according to region. Heatmap showing the similarity (Jaccard) between pairs of regions and within regions.*

Figure 4: *Correlations between community similarity with oceanic and environmental distance. Partial mantel correlations (R^2) showing the relative contribution of oceanic and environmental distance shaping the community structure for the different biological groups. The Mantel tests is based on Pearson's product moment correlation using 9999 permutations (see Table 2). Blue: oceanic distance; Green: environmental distance; Grey: Shared*

Figure 5: *Community similarity vs. ocean distance (a) and vs. environmental distance (b) for benthos sessile and mobile assemblages. Community similarity is fitted with a logarithmic decay model. Scatter plots depict data variability at each distance interval.*

Figure 6: *Tile diagram showing the Mantel correlations of the mobile and sessile community with environmental factors. The Mantel r values fall in the range of −1 to +1. Brighter cells with negative values in the plot indicate a negative correlation and redder cells with positive values a strong positive correlation. An r value of 0 indicates no correlation between community similarity and environmental variables. T: sea surface temperature (°C), T_range: SST range (°C), S: salinity, Chl-a: chlorophyll a (µl), PAR: Photosynthetic Available Radiation ($E.m^{-2}.day^{-1}$).*

Table 1. **Correlations between benthic community similarity with ocean distance and environmental factors**. Mantel correlations and Multiple Regression on distance Matrices (MRM) between species dissimilarity (Jaccard), log environmental distance, and log ocean distance between pairs of sampling sites; and Mantel partial correlations after controlling for the effects of either environmental or ocean distance, in statistically significant cases. N sites: number of biological sites considered at each group. The statistical significance of comparisons is assessed using Mantel and partial Mantel tests based on Pearson's product moment correlation using 9999 permutations. Environmental variables used (Bio-Oracle 2000-2014): SST, SST range, SSS, Chl, DAC, PAR. All cases significant (p-value = 0.05), except*.

Benthic groups	**Mantel correlation**			**Mantel partial correlation**		**MRM**
	N sites	**Ocean distance**	**Environmental distance**	**Ocean distance out environment**	**Env. out ocean distance**	**Ocean distance + environmental**
Sessile	42	0.79	0.67	0.44	0.07*	0.80
Mobile	36	0.87	0.76	0.68	0.12	0.88

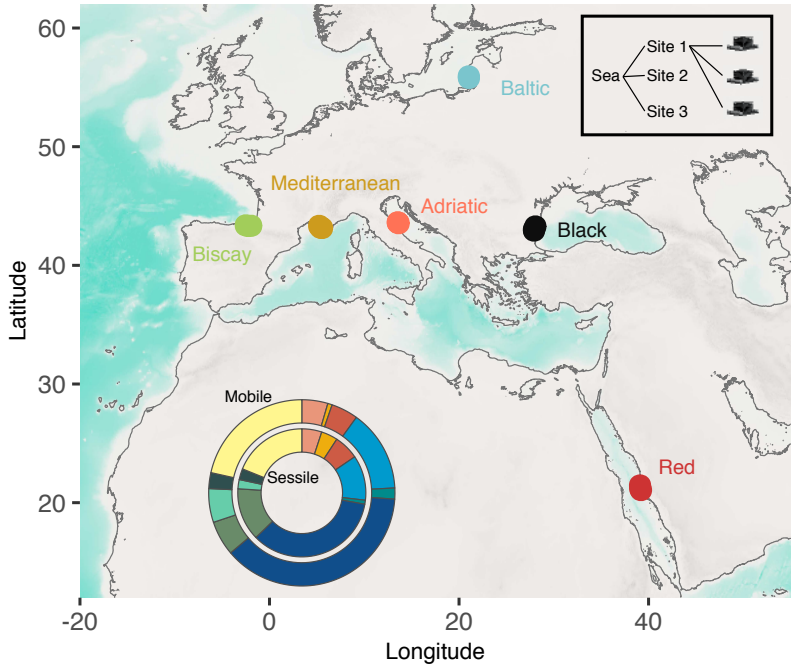
Table 2. **Correlations between benthic community similarity with ocean distance and environmental factors**. Mantel correlations and Multiple Regression on distance Matrices (MRM) between species dissimilarity (Jaccard), log environmental distance, and log ocean distance between pairs of sampling sites; and Mantel partial correlations after controlling for the

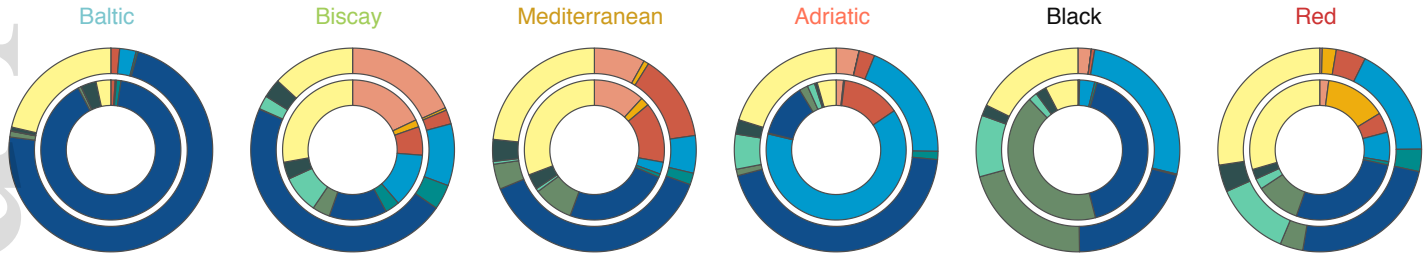
effects of either environmental or ocean distance, in statistically significant cases. N sites: number of biological sites considered at each biological group. The statistical significance of comparisons is assessed using Mantel and partial Mantel tests based on Pearson's product moment correlation using 9999 permutations. Environmental variables used (Bio-Oracle 2000-2014): SST, SST range, SSS, Chl-a, PAR. Significant cases in bold (p-value = 0.001).

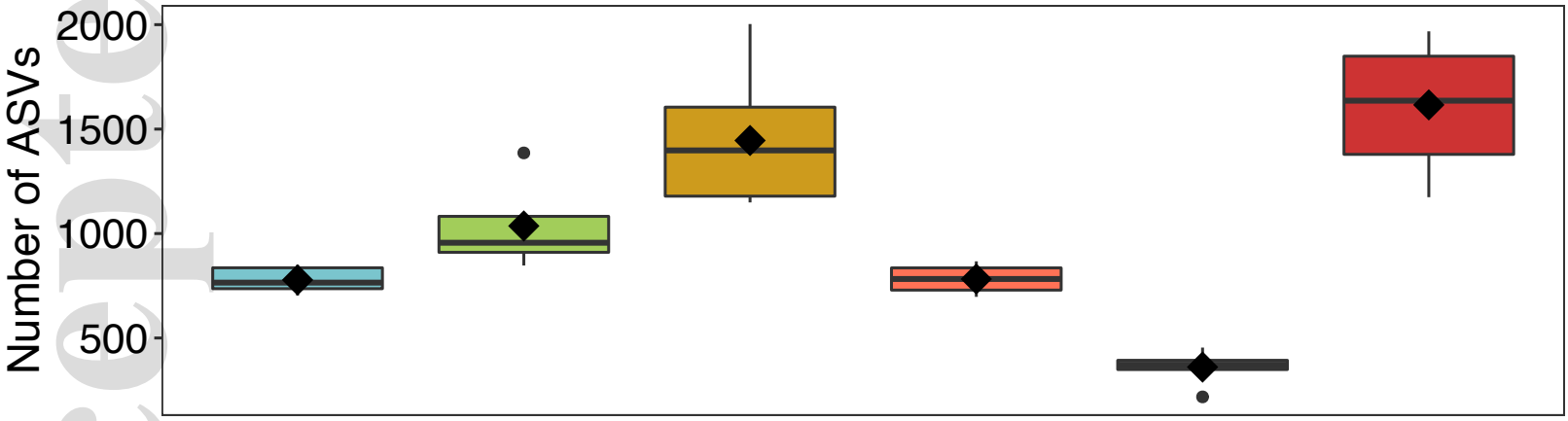
Phylum	**Mantel correlation (R^2)**			**Mantel partial correlation (R^2)**		**MRM (R^2)**
	N sites	**Ocean distance**	**Env. distance**	**Ocean distance out environment**	**Env. out ocean distance**	**Ocean distance + environmental**
Arthropoda	36	**0.77**	**0.76**	**0.33**	**0.27**	**0.79**
Chordata	36	**0.70**	**0.63**	**0.40**	0.06	**0.70**
Annelida	36	**0.80**	**0.77**	**0.43**	**0.22**	**0.81**
Mollusca	36	**0.69**	**0.70**	**0.24**	**0.26**	**0.72**
Echinodermata	36	**0.72**	**0.65**	**0.40**	**0.07**	**0.72**

Table 3. **Halving-distances derived from species similarity and oceanic distance which is calculated with a logarithmic decay model for sessile and mobile groups.** The logarithmic decay model shows the community similarity decline (slope) with the logarithm of oceanic distance. S_o is the initial community similarity at the "lowest" oceanic distance (100 km). The value of 100 km to obtain the S_o is imposed after analyzing the similarity-decay of each group along oceanic distances.

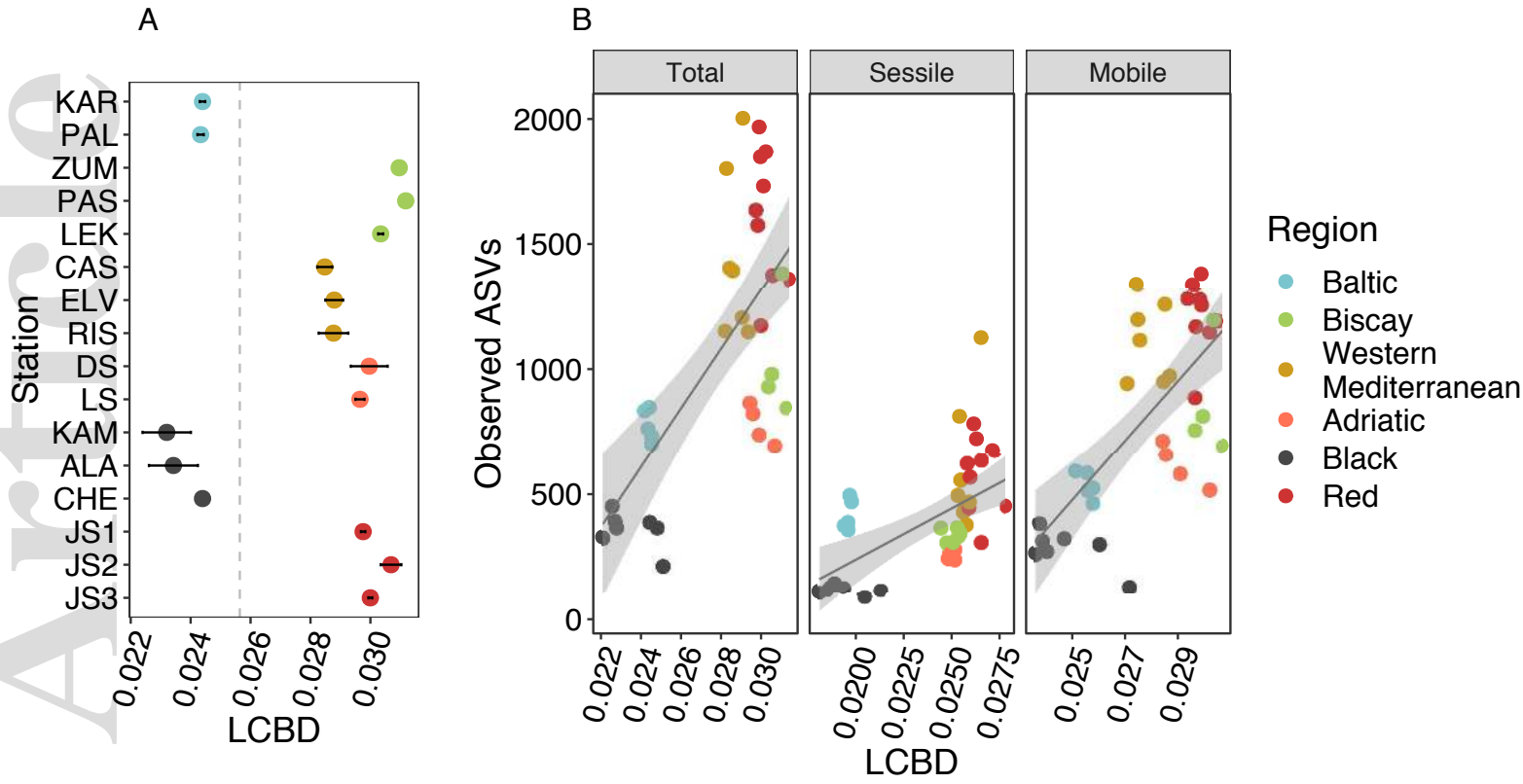
	Slope (c)	S_0= Initial similarity	Halving Distance (km)
Sessile	-0.0610	0.1476	251
Mobile	-0.0601	0.1481	261



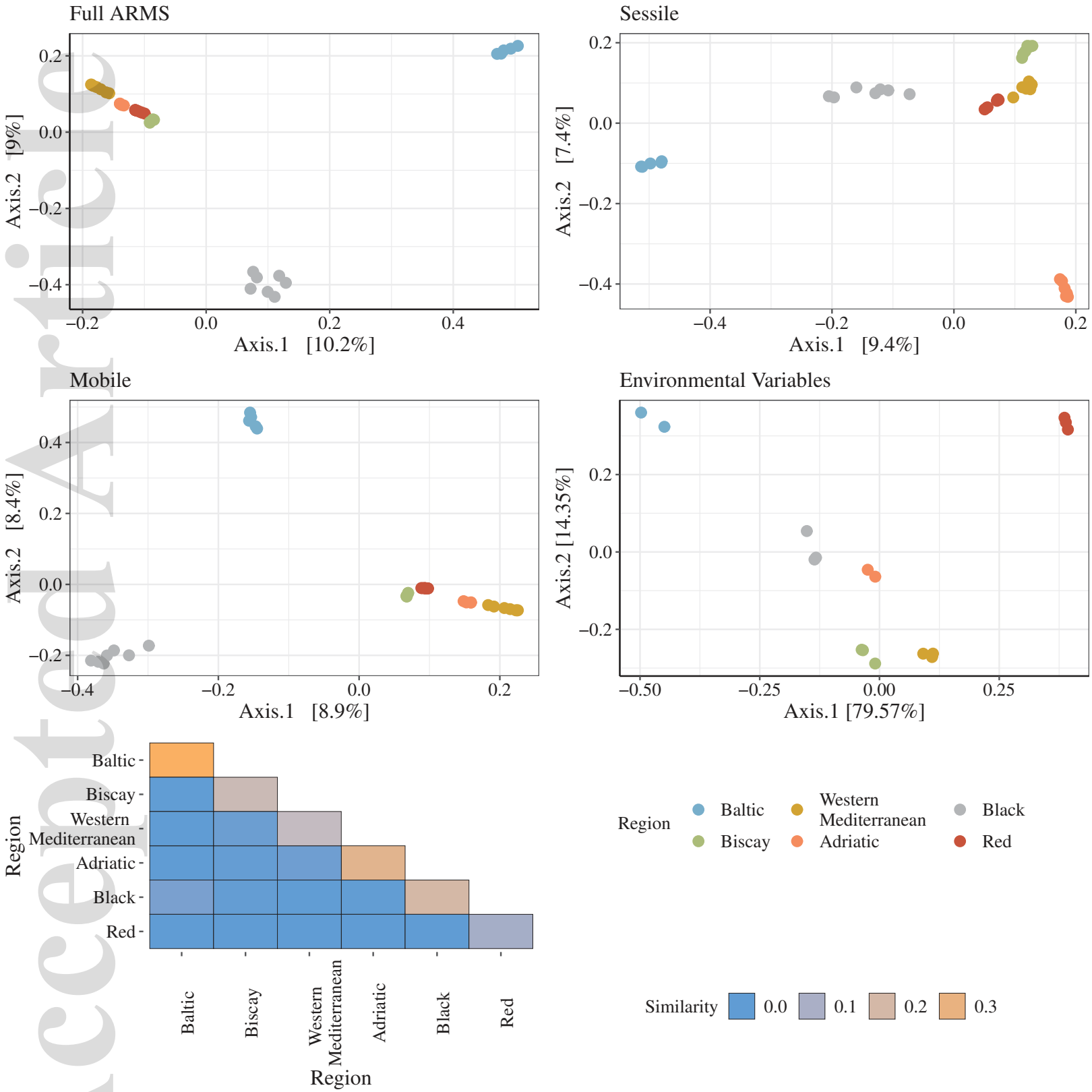




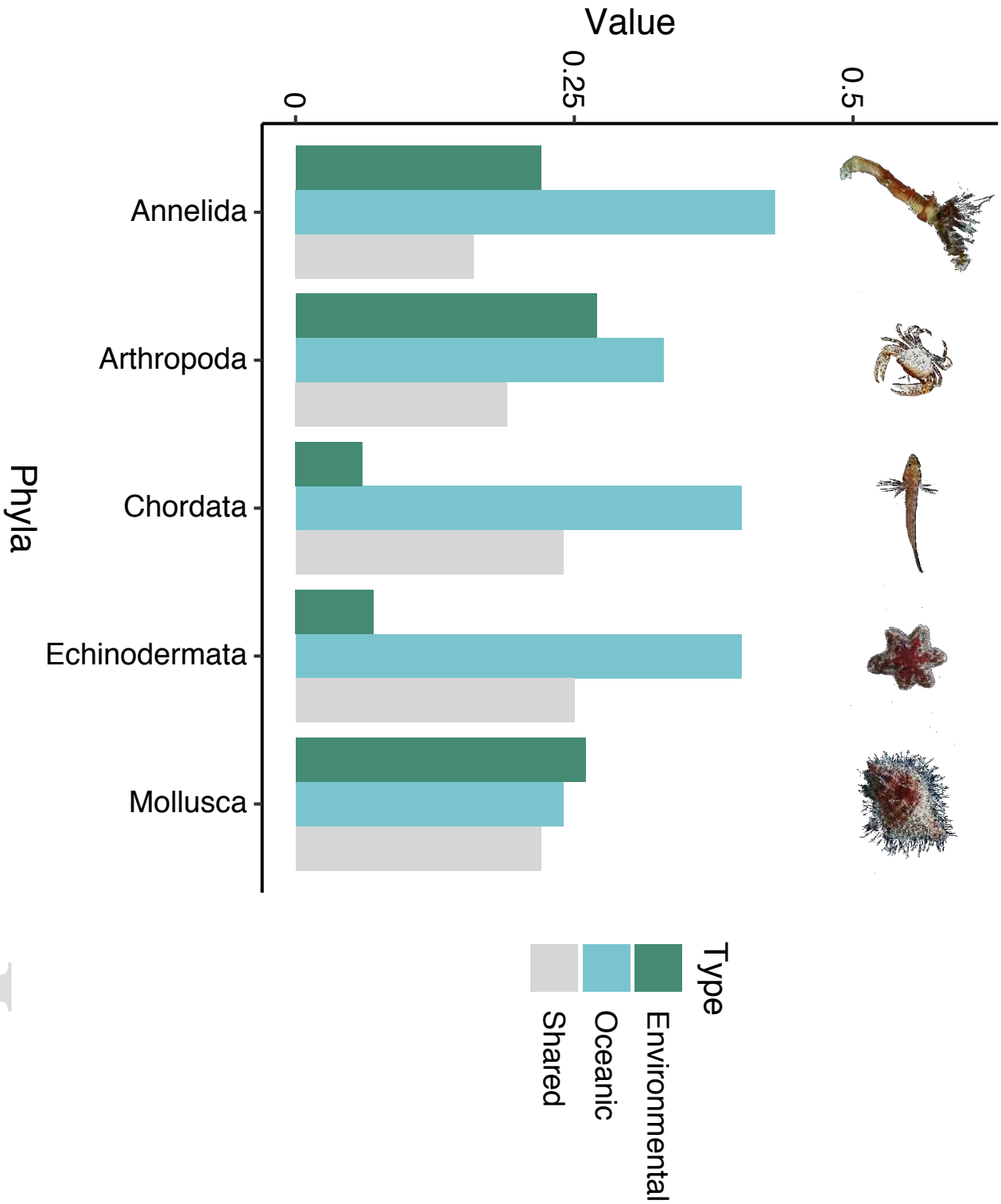
mec_15692_f1.eps



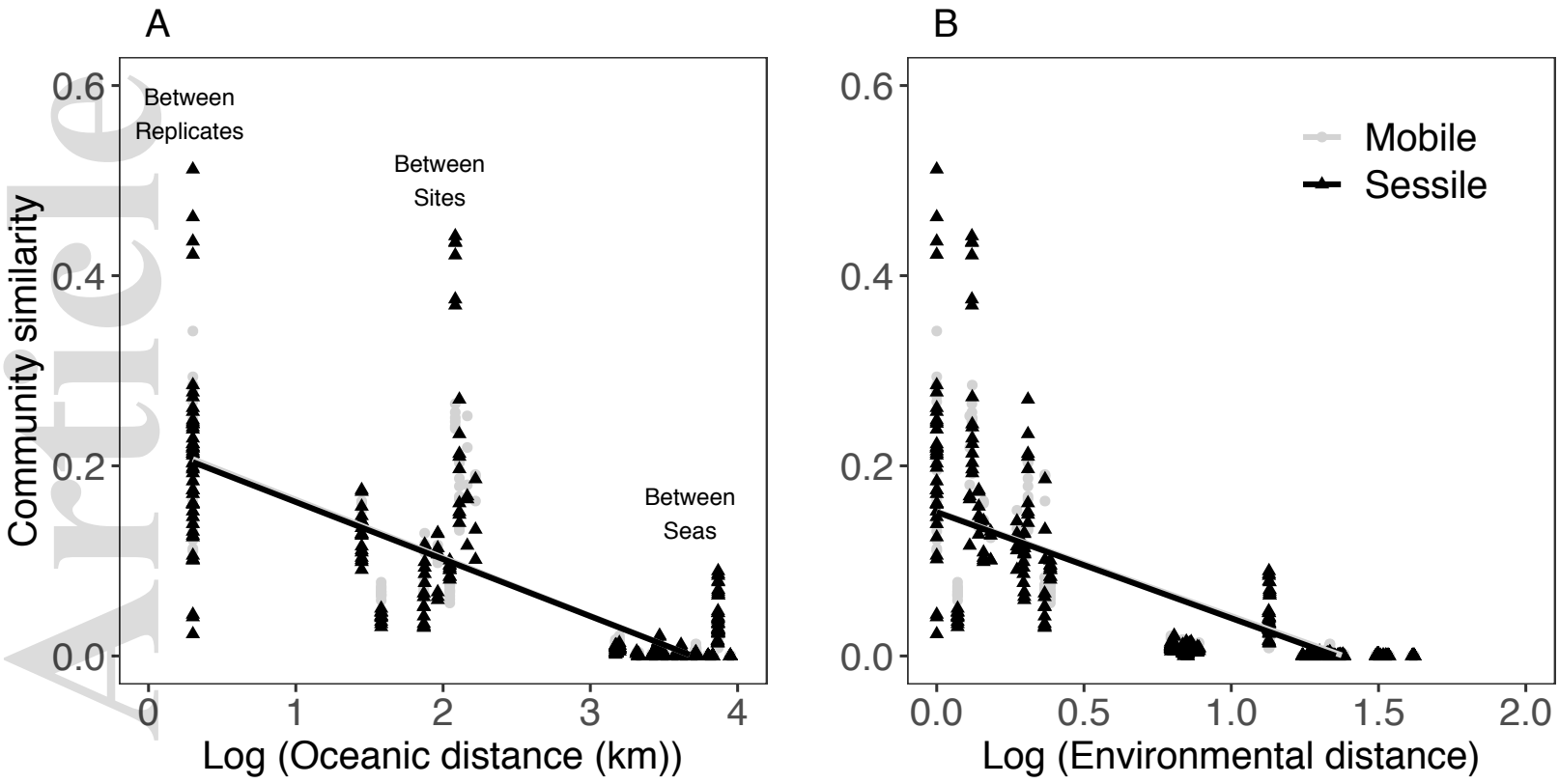
mec_15692_f2.eps



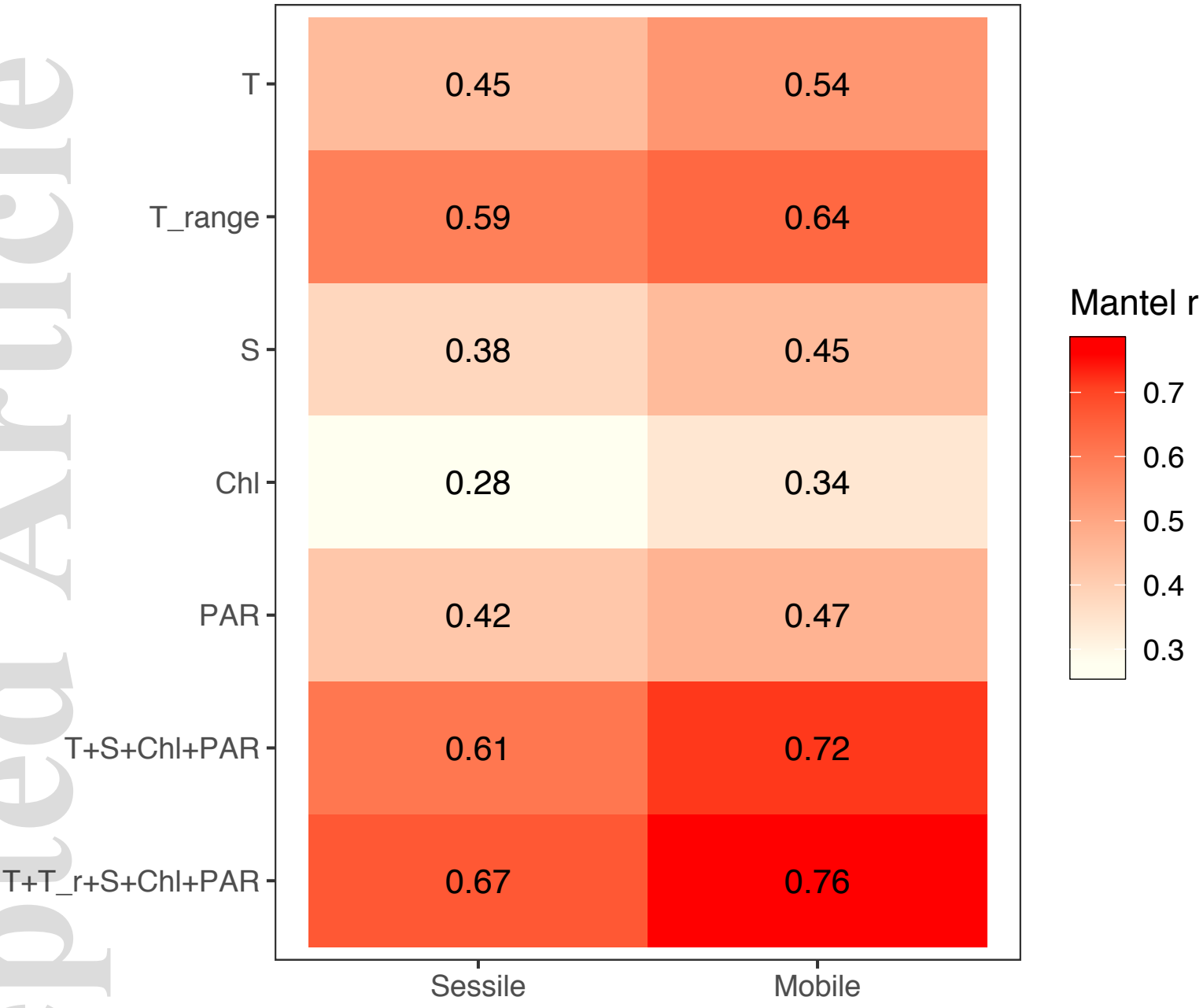
mec_15692_f3.eps



mec_15692_f4.eps



mec_15692_f5.eps



mec_15692_f6.eps