Universidade do Minho
Escola de Ciências

Bárbara Daniela da Rocha Leite

Monitoring Coastal Benthic Colonization of Artificial Substrates with the Support of DNA Metabarcoding Approaches

janeiro de 2021

Monitoring Coastal Benthic Colonization of Artificial Substrates with the Support of DNA Metabarcoding Approaches

Bárbara Daniela da Rocha Leite

UMinho | 2021

Universidade do Minho
Escola de Ciências

Bárbara Daniela da Rocha Leite

Monitoring coastal benthic colonization of artificial substrates with the support of DNA metabarcoding approaches

Ph.D. Thesis in Biology

Specialization in Integrated Management of the Sea

Work supervised by

Prof. Dr. Filipe José Oliveira Costa

Prof. Dr. Jesús Souza Troncoso

January 2021

Acknowledgements

I am grateful to the Portuguese Foundation for Science and Technology (FCT) for supporting my PhD grant PD/BD/127994/2016, hosted by the University of Minho (Portugal) and the University of Vigo (Spain), in the scope of the PhD programme in Marine Science Technology and Management (Do*Mar). I would also like to thank the project "The NextSea: Next generation monitoring of coastal ecosystems in a scenario of global change" (NORTE-01-0145-FEDER-000032), supported by NORTE 2020 (Norte Portugal Regional Operational Programme), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF), and the "Contrato-Programa" UIDB/04050/2020 funded by national funds through the FCT I.P. (Fundação para a Ciência e Tecnologia) for supporting this work.

I am grateful to the Centre of Molecular and Environmental Biology (CBMA, University of Minho) and CIM-UVIGO (ECIMAT Marine Station, University of Vigo) for receiving me and providing the facilities and resources for the development of my work.

To my supervisors, who made all this possible: Filipe Costa and Jesús Troncoso. They gave me the opportunity and the pleasure of being able to learn and evolve, with the freedom to work and make mistakes, but only in this way learning to make this thesis possible. To Filipe Costa, who always believed in me to work in the different projects we have already worked on, and have already been for some years. For the calm and confidence you have given me in these four years. To Jesús Troncoso, for the integration and the opportunity to integrate new projects, from where new ideas and projects emerged. To all stays in Vigo, due to the prompt availability of all the work conditions and services that I needed. I have no doubt that I had the best with me! Thank you!

To Conceição Egas and Guillermo Díaz-Agras for their availability to integrate the monitoring committee of the Annual Research Plan. For the cooperation and effort, through all the suggestions helped me to boost the work for better results.

To my sampling partner, Diego Carreira-Flores. A huge thanks for always understood my anxieties on the diving days. Thank you for being ready to collaborate, for the long hours of work and talking, and for the hospitality.

To Professor Pedro Gomes, for the help and camaraderie in sampling campaigns. For the conversations on the trips, and for the effort to endure the “melguinha”.

A special thanks to Pedro Vieira and Sofia Duarte for helping in thesis development. Still to Pedro, for the contribution in bioinformatics analysis and to help in peracarids identifications. To Marcos Teixeira for the help in the identification of polychaetes.

To my lab colleagues, to those who still are there and to those who have been there, one thanks for the good mood and friendship.

This PhD would not have been possible without all the foreign research groups and collaborations that supported me during my visits. I wish to thank to Marcos Rubal from CIIMAR for him help in sampling planning and campaigns. I wish to thank to Roberto Gómez and Enrique Poza from CIM-UVIGO (ECIMAT Marine Station) for their help during substrates deployment and sampling in Vigo. I acknowledge the support from Alberto Meisozo from Marine Biology Station of A Graña for him help on ARMS construction. Also from Marine Biology Station of A Graña (University of Santiago de Compostela), I wish to thank to Ramiro Tato and Guillermo Díaz-Agras for their help during substrates deployment and sampling in Ferrol.

A todos os meus amigos. Em especial às minhas Baleolas, Benedita, Jéssica, Marina e Susana, em que estes 4 anos foram a soma de mais bons momentos dentro de tudo o que temos vivido juntas, e que sejamos sempre assim: unidas, amigas e companheiras. À Bárbara e à Sofia, que sempre trouxeram momentos de alegria e ânimo durante estes anos. À Vera e ao Paulo, a família de Braga.

Aos amores da minha vida, Jorge e Maria. Os meus pais, que me deram a educação para lutar por aquilo em que acredito, e que sempre me deram o apoio e o amor para me tornar na pessoa que sou hoje. A toda a minha família, obrigada!

Ao meu marido, tudo! Meu amigo, meu companheiro, meu confidente. Obrigada pela paciência nestes tempos difíceis e diferentes, mas juntos somos mais fortes e contigo foi tudo mais fácil de superar. Depois de sobrevivermos a tudo isto....é para sempre!

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

Abstract The aptitude and potential of DNA metabarcoding for high-throughput monitoring of marine macrozoobenthos has been recently demonstrated. However, there are still significant challenges retarding the widespread implementation of DNA-based monitoring strategies in coastal ecosystems. In this thesis, we tested a DNA metabarcoding-based system, coupled with artificial substrates, for monitoring marine macrozoobenthic communities. To this end, we a) assembled a comprehensive reference library of DNA barcodes for marine macroinvertebrate species of Atlantic Iberia, and evaluated gaps in species coverage, b) investigated the efficiency of different marker loci (COI and 18S) and primer-pairs in macrozoobenthos detection, c) compared the performance of DNA metabarcoding and morphology in species detection, d) compared the impact of artificial substrates, made of different materials (slate, PVC and granite) and shapes (ARMS and ASMS), on the patterns of colonization of macrozoobenthos, and e) investigated the temporal and regional variation in macrozoobentos from artificial substrates deployed in NW Atlantic Iberia. Through the compilation of a comprehensive reference library, we recorded a large portion of taxa pending barcode coverage (63%) and signaled a high proportion of species with significant intraspecific divergence (16%). The high complementary of species detection among primers and markers (maximum 13% overlap) indicated that no single marker or primer can provide a complete diagnosis of the macrozoobenthos diversity. Until more extensive monitoring data is available, DNA metabarcoding and morphology should be used in combination, to avoid missing relevant taxa and obtain abundance estimates. Unlike substrate materials, substrate shape strongly affected the colonization of species. Taxonomic diversity varied between substrates, especially when comparing ARMS and ASMS (maximum overlap 30%). Important fractions of diversity may be overlooked if only one substrate is used for monitoring. Substrate deployment periods also displayed a strong influence in the colonization of macrozoobenthos, signaling the importance of this factor for the monitoring design. We conclude that DNA metabarcoding combined with artificial substrates (especially through the combination of ARMS and ASMS) have great potential to be used in comprehensive coastal biomonitoring programs targeting macroinvertebrate biodiversity.

Keywords Artificial substrates; DNA metabarcoding; Marine macrozoobenthic communities; Reference library; Short- and long-term monitoring.

Resumo A capacidade e o potencial do *DNA metabarcoding* para a monitorização de alto rendimento do macrozoobentos marinho têm sido demonstrados em estudos relativamente recentes. Contudo, existem ainda desafios significativos que têem retardado a sua implementação generalizada na monitorização de ecossistemas costeiros. Nesta tese, testamos uma abordagem de monitorização de comunidades macrozoobentónicas marinhas que conjuga *DNA metabarcoding* com substratos artificiais. Para esse feito, a) criámos uma biblioteca de referência de *DNA barcodes* para espécies macrozoobentónicas marinhas da Ibéria Atlântica, e avaliámos as lacunas na cobertura de espécies, b) investigámos a eficiência de diferentes marcadores moleculares (COI e 18S) e de *primers* na deteção de macrozoobentos, c) comparámos o desempenho entre morfologia e *DNA metabarcoding* na deteção de espécies, d) comparámos o impacto do material (ardósia, PVC e granito) e da forma (ARMS e ASMS) dos substratos artificiais nos padrões de colonização de macrozoobentos, e e) investigámos variações temporais e regionais na colonização de macrozoobentos em substratos artificiais implantados no noroeste Ibero-Atlântico. Através da compilação da biblioteca de referência registou-se uma porção elevada de *taxa* sem *DNA barcode* (63%) e sinalizou-se uma grande proporção de espécies com divergência intraespecífica significativa (16%). A elevada complementaridade entre *primers* e marcadores na deteção de espécies (máximo 13% sobreposição), indicou que nenhum marcador ou *primer* consegue fornecer individualmente um diagnóstico completo da diversidade macrozoobentónica. Enquanto não estiveram disponíveis dados de monitorização mais sistematizados, tanto o *DNA metabarcoding* como a morfologia deverão ser utilizados, de modo a não discurar de *taxa* relevantes, e a obter estimativas de abundância. Contrariamente ao material, a forma dos substratos afetou fortemente a colonização de espécies. A diversidade taxonómica variou entre todos os substratos, especialmente ao comparar ARMS e ASMS (máximo 30% sobreposição). Frações importantes de diversidade poderão ser subestimadas se for apenas usado um substrato para monitorização. Os períodos de submersão dos substratos influenciaram fortemente na colonização de macrozoobenthos, evidenciando a importância deste fator no desenho da monitorização. Concluímos que o *DNA metabarcoding* conjugado com substratos artificiais (nomeadamente ARMS e ASMS) tem um potencial elevado para aplicação em programas de biomonitorização costeira que visem a biodiversidade de macroinvertebrados.

Palavras-chave Biblioteca de referência; Comunidades macrozoobentónicas marinhas; *DNA metabarcoding*; Monitorização de curta e longa duração; Substratos artificiais.

Table of contents

Figures index xii

Tables index xvi

Abbreviations xvii

Chapter 1 General introduction 1

1.1 Marine biodiversity 2

1.2 Coastal macrozoobenthic communities 4

1.3 Coastal macroinvertebrate biomonitoring 6

1.3.1 Challenges of monitoring macrozoobenthic communities 6

1.3.2 Taxonomic identification methodologies 8

1.4 DNA-based approaches for macrozoobenthic biodiversity research and monitoring 9

1.4.1 DNA metabarcoding developed studies 9

1.4.2 Challenges of DNA metabarcoding approaches in coastal macrozoobenthic communities ... 11

1.5 Aims and structure of the thesis 13

References 18

Chapter 2 Gap-analysis and annotated reference library for supporting macroinvertebrate metabarcoding in Atlantic Iberia 34

Abstract 35

2.1 Introduction 36

2.2 Material and methods 38

2.2.1 Study area and checklist 38

2.2.2 Data mining and BOLD dataset creation 38

2.2.3 Data processing and analyses 39

2.3. Results 41

2.4 Discussion 44

2.5 Conclusions 49

References 50

Chapter 3 Comparing species detection success between molecular markers and primers in DNA metabarcoding of coastal macroinvertebrate communities 59

Abstract 60

3.1 Introduction 61

3.2 Materials and methods 64

3.2.1 Sampling design 64

3.2.2 DNA metabarcoding-based taxonomic identification 65

3.2.3 Community analyses 68

3.2.4 Gap-analysis and species discrimination ability 69

3.3 Results 70

3.3.1 COI-5P primer-pairs testing 70

3.3.2 Effect of marker and primer choice on species detection 71

3.3.3 Availability of reference sequences on public databases 75

3.4 Discussion 77

References 82

Chapter 4 Coastal macrozoobenthos monitoring: impact of artificial substrate material and comparison of morphology vs DNA metabarcoding-based taxonomic profiling 90

Abstract 91

4.1 Introduction 92

4.2 Materials and methods 94

4.2.1 Sampling design 94

4.2.2 Morphology-based taxonomic identification 95

4.2.4 DNA metabarcoding-based taxonomic identification 95

4.2.5 Community statistical analyses 97

4.3 Results 99

4.3.1 Morphology-based macrozoobenthic taxonomic assignments 100

4.3.2 DNA metabarcoding-based macrozoobenthic taxonomic assignments 103
4.3.3 Morphology and DNA metabarcoding comparison 104
4.3.4. Ecological successional patterns: changes in community composition 106
4.4 Discussion 112
4.4.1 Comparison of species detection success employing morphology or DNA-based approaches 113
4.4.2 Substrates effects and seasonal succession patterns on macrozoobenthic community 115
4.5 Conclusion 117
References 118

Chapter 5 ARMS embracing ASMS to complement DNA metabarcoding-based monitoring of macrozoobenthos in NW Iberian coast 125
Abstract 126
5.1 Introduction 127
5.2 Materials and methods 129
5.2.1 Study area 129
5.2.2 Sampling design 130
5.2.3 ARMS and ASMS recovery and processing 130
5.2.4 DNA metabarcoding 132
5.2.5 Data processing 133
5.2.6 Statistical analyses 134
5.3 Results 135
5.4 Discussion 146
5.5 Conclusion 151
References 152

Chapter 6 Concluding remarks 159
References 165

Annexes 167
Annexes of Chapter 3 168
Annexes of Chapter 4 180
Annexes of Chapter 5 195

Figures index

Figure 1.1. Representative scheme of the structure of the thesis, with the connections between chapters. 17

Figure 2.1. A. DNA barcode coverage for marine macroinvertebrate species occurring in Atlantic Iberia. Barcode coverage with at least one COI-5P sequence per species (green bar). Numbers on the right side of each bar refer to the total number of sequences. **B.** Partioning of the geographic origin of specimens available in the reference library: samples in Atlantic Iberia, outside of Iberian Peninsula and uncertain geographical information (unknown).....42

Figure 2.2. Number of BINs according to taxonomic congruence annotations (concordant, multiple or discordant) for each taxonomic group on the Atlantic Iberia reference dataset. 44

Figure 3.1. Number of marine macroinvertebrate species detected in each substrate and sampling time by each of the four primer-pairs used in the first screening of primer performance.....70

Figure 3.2. Shared and unique marine macroinvertebrate species detected (A) by the four COI-5P primer-pairs (I – mlCOIintF/LoboR1; II – mlCOIintF/jgHCO2198; III – LCO1490/III_C_R; IV – LoboF1/230R), and (B) by the two primer-pairs targeting the 3' end (mlCOIintF/LoboR1 and mlCOIintF/jgHCO2198). 71

Figure 3.3. Taxonomic profile of the marine macroinvertebrate species detected in the substrates by each primer-pair. 72

Figure 3.4. Number of species detected, partitioned among the 9 phyla recorded and by primer-pair employed. a and b indicate significant differences between the primers ($p<0.05$). 73

Figure 3.5. Taxonomic composition of marine macroinvertebrate communities for each primer-pair in each substrate type (A - Slate, B - PVC and C - Granite) and sampling time (3, 7, 10 and 15 months). . 74

Figure 3.6. Partitioning of the marine macroinvertebrate species detection for (A) both marker loci and (B) primer-pair, in the three substrate types and among all sampling times. 75

Figure 3.7. Availability of reference sequences of COI-5P and 18S V4 for each taxonomic group of marine macroinvertebrate species detected with the three primer-pairs from COI-5P and 18S genes. Barcode coverage with at least one sequence per species (black bar). 76

Figure 4.1. Map of the Iberian Peninsula showing the sampling site (black star: Toralla Island, NW Iberian Peninsula) and the three artificial substrates (10 x 10 cm) used for marine macrozoobenthos colonization. A – Slate; B – PVC; C – Granite. 94

Figure 4.2. Photos of the artificial substrates (slate, PVC and granite) collected after 3, 7, 10 and 15 months of deployment. 99

Figure 4.3. Number of detected taxa using morphology in each substrate type and sampling time (3, 7, 10 and 15 months). Total: total number of detected taxa in all substrates combined for each sampling time. 102

Figure 4.4. Number of detected taxa using DNA metabarcoding in each substrate type and sampling time (3, 7, 10 and 15 months). Total: total number of detected taxa in all substrates combined for each sampling time. 103

Figure 4.5. Comparison between morphological and DNA metabarcoding species-level identifications in three substrates types (Slate, PVC and Granite) with sampling time data merged. The upper bar chart shows the distribution of the total number of species per phylum obtained either by morphology or DNA metabarcoding, in each substrate type. The Venn diagrams in the lower part shows the proportion of species detected exclusively by morphology (blue), exclusively by DNA metabarcoding (orange), and shared by both approaches (overlapping circles), for each substrate type and for all combined substrates. Differences were detected between both approaches (cophenetic correlation coefficient of the Bray-Curtis similarity = 0.51). 105

Figure 4.6. Non-metric multidimensional scaling (nMDS), based on Bray-Curtis similarity index, of marine macrozoobenthic species detected in the artificial substrates (Slate, PVC and Granite) deployed in Toralla Island. Colors group species identified by different methodologies: morphology (blue) and DNA metabarcoding (orange). Sampling times: 3M – 3 months; 7M – 7 months; 10M – 10 months; 15M – 15 months..... 106

Figure 4.7. Total detected taxa for each artificial substrate (slate, PVC and granite) and for the total species detected between the three substrates (Total), among sampling times (3, 7, 10 and 15 months). 107

Figure 4.8. Proportion of the species detected exclusively and shared between the three substrates (Slate, PVC and Granite) for each sampling time (3, 7, 10 and 15 months), and for the combination of the total species detected in each substrate (global). 108
Figure 4.9. Heatmap of marine macrozoobenthic species detected in each substrate and sampling time. Substrates are grouped by colors corresponding to months of deployment; 3M – 3 months; 7M – 7 months; 10M – 10 months; 15M – 15 months. Green represents species occurrence. 109
Figure 4.10. Average taxonomic distinctiveness (AvTD) funnel plot of the marine macrozoobenthic communities recorded in artificial substrates (slate, PVC and granite) among sampling times. 3M – 3 months; 7M – 7 months; 10M – 10 months; 15M – 15 months. 111
Figure 4.11. Non-metric multidimensional scaling (nMDS), based on Bray-Curtis similarity index, of the total species detected in the artificial substrates (Slate, PVC and Granite) among sampling times. Sampling times: 3M – 3 months (yellow); 7M – 7 months (orange); 10M – 10 months (purple); 15M – 15 months (blue). 112
Figure 5.1. Study sites (black star): **A** – Bajo Tofiño, Ría de Vigo, Spain; **B** – San Cristovo, Ría de Ferrol, Spain.....129
Figure 5.2. Artificial substrates used for marine macrozoobenthic colonization. A – ARMS; B – ASMS. 130
Figure 5.3. ARMS recovery procedure. 131
Figure 5.4. ASMS recovery procedure. 131
Figure 5.5. Sampled ARMS face plates collected after 6, 9 and 12 months of deployment at Bajo Tofiño - Ría de Vigo, and San Cristovo - Ría de Ferrol. A – Plate 1 top; B – Plate 5 bottom; C – Plate 9 top; D – Plate 9 bottom. 135
Figure 5.6. ASMS collected after 6, 9 and 12 months of deployment at Bajo Tofiño - Ría de Vigo and San Cristovo - Ría de Ferrol. 136
Figure 5.7. Taxonomic distribution of the total species detected in each substrate: ARMS and ASMS. 137
Figure 5.8. ARMS and ASMS substrates from two sampling sites (A) Ría de Ferrol (RF) and (B) Ría de Vigo (RV). Bar charts represents the abundance and taxonomic distribution of species detected on each substrate among sampling times (6M – 6 months; 9M – 9 months; 12M – 12 months). 138

Figure 5.9. Partitioning of the marine macrozoobenthic species detected exclusively by ARMS (grey), exclusively by ASMS (blue), and shared by both substrates (overlapping circles) for all sampling times in (A) Ría de Vigo and (B) Ría de Ferrol; and (C) the combined proportion of species detected by both substrates in each sampling site. 140

Figure 5.10. Heat-map of marine macrozoobenthic species most accounting for dissimilarities between substrates, after SIMPER analysis. RV - Ría de Vigo; RF - Ría de Ferrol. 6M – 6 months; 9M – 9 months; 12M – 12 months. Blue represents species detection. 141

Figure 5.11. Non-metric multidimensional scaling (nMDS), based on Bray-Curtis similarity index, of the species detected in both artificial substrates for mobile and sessile fauna fractions separately. Samples combine all species detected by both substrates on each sampling site (RV – Ría de Vigo; RF – Ría de Ferrol) among sampling times (6M – 6 months, 9M – 9 months and 12M – 12 months)..... 143

Figure 5.12. Fluctuations in presence/absence of taxa for most abundant taxonomic groups within ARMS and ASMS from Ría de Vigo (RV) and Ría de Ferrol (RF), and for all taxa detected in each Ría divided by sampling seasons (6 months – winter, 9 months – spring, 12 months – summer)..... 144

Figure 5.13. Confidence funnel plot of the average taxonomic distinctiveness (AvTD) of the marine macrozoobenthic communities recorded in ARMS and ASMS from Ría de Vigo (RV) and Ría de Ferrol (RF), and across sampling times. Thin line represents theoretical mean (95% confidence interval). 6M – 6 months (blue); 9M – 9 months (yellow); 12M – 12 months (pink)..... 145

Figure 5.14. Non-metric multidimensional scaling (nMDS), based on Bray-Curtis similarity index, of the species detected in ARMS (■ - green circles) and ASMS (● - red circles). Samples combine the total species detected on each sampling site (RV – Ría de Vigo and RF – Ría de Ferrol) among sampling times (6 months - blue, 9 months - yellow and 12 months - pink)..... 146

Tables index

Table 2.1. Number of COI-5P sequences generated under this study and number of COI-5P sequences retrieved from BOLD, compiled in the dataset DS-GAIMARIN for each target taxa, with the associated BOLD project code. 39

Table 3.1. Primer-pairs used to test the efficiency of COI-5P to amplify and assess marine macroinvertebrate species. F – forward; R – reverse; bp – base pairs.....66

Table 3.2. Primer-pairs and respective thermal cycling conditions used in this study to amplify marine macroinvertebrate communities. F – forward; R – reverse; bp – base pairs. 67

Table 4.1. Diversity analysis for the univariate indices of the assemblages detected in each substrate type across sampling times, using morphology. *S* – species richness. *N* – species abundance. *d* – Margalef species richness index. *J'* – Pielou's evenness index. *H'* – Shannon-Wiener index. 3M – 3 months; 7M – 7 months; 10M – 10 months; 15M – 15 months. 1, 2 and 3 corresponds to the replicate number.....101

Table 4.2. SIMPER analysis showing species-level differences in community composition across all sampling times. 3M – 3 months; 7M – 7 months; 10M – 10 months; 15M – 15 months. 110

Table 5.1. Primer-pairs and respective thermal cycling conditions used to amplify marine macrozoobenthic communities. F – Forward; R – Reverse; bp – Base pairs.....133

Table 5.2. SIMPER analysis showing species-level differences in community composition across substrates. 138

Table 5.3. SIMPER analysis showing species-level differences in community composition among sampling times and between substrates. 6M – 6 months; 9M – 9months; 12M – 12 months..... 142

Abbreviations

^{0}C	Celsius degree
µL	Microliter
µm	Micrometre
µM	Micromolar
18S rRNA	18S Ribossomal Ribonucleic Acid
ANOSIM	One-way Analyses of Similarities
ANOVA	Analysis of Variance
ARMS	Autonomous Reef Monitoring Structures
ASMS	Artificial Seaweeds Monitoring System
AvTD	Average Taxonomic Distinctiveness
BIN	Barcode Index Number
BOLD	Barcode of Life Data Systems
bp	Base pair
CLUSTER	Hierarchical clustering
COI	Cytochrome c oxidase subunit I
COI-5P	5' end of the cytochrome c oxidase subunit I gene corresponding to barcode region
d	Margalef species Richness Index
DNA	Deoxyribonucleic Acid
EDTA	Ethylenediaminetetraacetic Acid
e.g.	*exempli gratia*
et al.	*Et alia*
F	Forward reads
H'	Shannon-Wiener Diversity Index
HTS	High-throughput Sequencing
i.e.	*id est*
J'	Pielou's Evenness Index

K2P	Kimura-2-parameter
M	Month
mBrave	Multiplex Barcode Research and Visualization Environment
min	Minutes
mM	Milimolar
MSFD	Marine Strategy Framework Directive
mtDNA	Mitochondrial DNA
N	Total abundance
NaCl	Sodium Chloride
ng	Nanogram
NIS	Non-indigenous species
NJ	Neighbor Joining
nMDS	Non-metric Multidimensional Scaling
NW	Northwest
OTU	Operational Taxonomic Unit
p	p-Value
PAST	Plaeontological Statistics Software
PCR	Polymerase Chain Reaction
PERMANOVA	Permutational Multivariate Analysis
pH	Potential of hydrogen
PRIMER-E	Plymouth Routines in Multivariate Ecological Research
PVC	Polyvinyl Chloride
R	Reverse reads
rpm	Rotations per minute
s	Second
S	Species richness
SDS	Sodium Dodecyl Sulphate
SIMPER	Similarity Percentages Test

TD	Taxonomic Diversity
Tris-HCL	Trizma Hydrochloride
UPGMA	Unweighted Pair Group Method with Arithmetic Mean
WFD	Water Framework Directive
WoRMS	World Register of Marine Species

Chapter 1

General introduction

1.1 Marine biodiversity

The marine realm currently accounts for some of the most threatened ecosystems on Earth in spite of their relevance for human well-being. Coastal habitats comprise approximately 15% of the world's total area (4% in land and 11% of the earth's oceans), interconnecting coastal areas to the continental shelf, down to 50 meters depth (Barbier, 2017). However, marine biodiversity remains largely uncharacterized (Mora et al., 2011; Appeltans et al., 2012). From the 2 million species described, and the estimation of 10 to 100 million of possible extant species (Roskov et al., 2018), only about 10% are known marine species (Gage, 2001). We are far from having a comprehensive list of marine species and their coverage is still largely unbalanced across different taxonomic groups (Bouchet, 2006).

Coastal ecosystems are composed by the most productive and diverse communities with a wide range of taxa (Poore et al., 1993), with high importance to primary productivity (Solan et al., 2004). This privileged environment present variations in their species composition depending on seasonality and follow a process of colonization. Hydrodynamics, predation, light, temperature and salinity are ecological factors that account for macrozoobenthos distribution (Gage, 2001; Veiga et al., 2016).

In water ecosystems, species dispersal is wide-ranging without physical barriers, which allow for genetic connectivity over huge expanses. Climatic conditions fluctuate considerably between coastal areas, and the existence of zones which have never been connected to the mainland (e.g. Macaronesian archipelagos) resulting in ecosystems with limited dispersal capacity and *in situ* diversification (Vieira et al., 2019). Relatively to the patterns of colonization of benthic species, they are mainly explained by the high structural complexity of benthic ecosystems, which provides new, sometimes small-sized, colonizers with a shelter against predators (Hereu et al., 2005; Jenkins et al., 2009). Thus, habitat complexity is an important factor in marine ecosystems, since it directly affects patterns of short-term colonization (García-Sanz et al., 2012).

Comprehensive marine biodiversity assessments are essential to recognize environmental changes and track responses to efficiently conserve marine ecosystems. Human activities have been changing the climate and destabilizing coastal ecosystems, which have been highly exploited, lost and degraded (Lau et al., 2019).

In the last decades, the increase of human population and exploitation of natural resources in coastal areas, whether for tourism or for commercial activities, led to changes in ecosystems. Additionally, given the current loss of coastal area, the proliferation of artificial structures, especially for coastal defense, has been increasing (Firth et al., 2016). Another growing problem in marine environments is the persistent overfishing, which induces loss of biodiversity and overexploitation of natural resources (Jackson et al., 2001). In fact, the loss of fisheries has been associated to the decline of water quality, increasing the occurrence of algal blooms and offshore pollution (Barbier, 2017). Additionally, loss of filtration capacity associated to the harvest of fishery resources (e.g. oysters or sea cucumbers) intensifies oxygen depletion and high levels of nitrogen (Purcell et al., 2013). These synergistically effects of multiple human-induced stresses (e.g. habitat destruction, pollution, eutrophication, increasing of nutrient loading and sedimentation) contributed to the degradation of coastal ecosystems, leading to negative impacts and changes on ecological processes occurring there (Solan et al., 2004; Rabalais et al., 2009). These changes have effects on many marine species, especially in their metabolism, development, growth and reproduction, potentially affecting food webs (Clements et al., 2018; Gallego et al., 2020). This situation is particularly alarming for macrobenthic communities, due to their sedentary activity. The introduction of new species in the habitats (i.e. non-indigenous species – NIS), which can rapidly spread and become invasive and out-compete with native species, can led to critical ecological changes and possible loss of native species and their ecosystem services provided (Molnar et al., 2008; Simberloff et al., 2013).

Rigorous experimental frameworks to provide useful ecological advice and guidance, based on consequent ecological changes in marine coastal ecosystems (i.e. eco-engineering) is essential for marine coastal management (Chapman et al., 2018). However, due to rare and small species (Albano et al., 2011), sometimes excluded from marine biodiversity assessments (Roberts et al., 2002; Tittensor et al., 2010), and the huge diversity of taxa (Templado et al., 2010), marine macrozoobenthic communities are challenging to assess and monitor.

1.2 Coastal macrozoobenthic communities

Macrozoobenthic species are benthic organisms mainly retained in a 0.5-1.0 mm mesh net (McLachlan and Defeo, 2017), accounted for the most taxonomically diverse category of benthos (Gage, 2001). The current recognized macrozoobenthic phyla are Annelida, Arthropoda (namely Crustacea), Brachiopoda, Bryozoa, Chordata (namely Tunicata), Cnidaria, Echinodermata, Mollusca, Nematoda, Nemertea, Platyhelminthes, Porifera and Sipuncula, where Polychaeta, Mollusca and Crustacea are the predominant taxa (McLachlan and Defeo, 2017).

These organisms may be attached (i.e. sessile macrobenthos, e.g. ascidieans or barnacles) or have characteristics to move over hard substrates (i.e. mobile macrobenthos, e.g. amphipods) (Gage, 2001; Chapman, 2003). Relatively to feeding habits they could be: a) suspensivores, feeding on particles suspended on water (e.g. sponges or ophiurids), b) detritivores, feeding on deposited particles settling from the water column (e.g. gastropods), c) herbivores, which scrap and ingest algae (e.g. sea urchins), or d) predators, which feed other organisms (e.g. crabs) (Gage, 2001). Additionally, some benthic species have different dynamics in food webs and are omnivorous (e.g. isopods as omnivorous scavangers; Lalli and Parsons, 1997). For colonization processes, macrobenthos usually arrive in a larvae phase or emigrate as sub-adults via crawling, drifting or rafting through the water column and disperse for some distance from their parental populations (Chapman, 2002).

Macrozoobenthic assemblages are complex and variable over space and time, which difficult the understanding of the ecological processes and the abundance and distribution of species (Rubal et al., 2014; Veiga et al., 2014). However, many species are recognized as ecosystem engineers, due to their ability to increase habitat complexity and to protect from negative effects of abiotic and biotic factors (Rubal et al., 2018). For example, peracarids are contributors to benthic production and an important source of food for other organisms (Moreira et al., 2008; Guerra-García et al., 2011). Other groups, such as Porifera, are ecologically and commercially important (e.g. pharmaceutical industries) (Bucklin et al., 2011).

The structure and function of marine ecosystems provides a wide range of benefits to human societies, i.e. ecosystem services. The coastal zoobenthic communities produce invaluable goods (i.e. products obtained from fishery) and services (e.g. recreation, storm protection or pollution control), which have a direct impact on human well-being (Barbier, 2017). However, the limited number of studies developed on marine water

underestimate the ecosystem services supplied (Barbier et al., 2012). The recognition and full capture of the biodiversity present on coastal habitats is essential to estimate the services that they could supply. Additionally, macroinvertebrate species are responsible for important habitat structure and processes (e.g. filtration or decomposition). For example, polychaete tubes lawns in coastal habitats act as fast-growing opportunistic species for the recolonization of disturbed substrates being the key for further benthic succession (Friedrichs et al., 2000). In macrozoobenthic communities, mobile and sessile organisms display different roles to supply ecosystem services. While sessile fauna, due to their ability to settle and colonize substrate, have importance, for example, in coastal erosion protection, mobile organisms are able to supply services related to pollination, seed dispersal and decomposition (Lundberg and Moberg, 2003).

The ecological status of a marine ecosystem may affect the ability of the ecosystem to provide services, and consequently the human survival and well-being (Liquete et al., 2013; Culhane et al., 2018). The rapid rate of how marine ecosystems have been losing biodiversity and disappearing in a worldwide scale (e.g. coral reefs; Grabowski et al., 2012) highlight the need to assess and evaluate their benefits, and to monitor these ecosystems to improve coastal management and policy.

In response to the increasing impacts on water ecosystems, worldwide efforts have been focused on developing and applying metrics to assess the conditions of aquatic ecosystems, to protect, restore and achieve a good status (Borja et al., 2010; Hering et al., 2010). The Water Framework Directive (WFD; Directive 2000/60/EC) was implemented to protect and improve transitional and coastal European waters, based on the implementation of ecological quality elements to compare the structure of these ecosystems, and achieve a "good ecological status" (Borja et al., 2009). The Marine Strategy Framework Directive (MSFD; Directive 2008/56/EC) was developed to effectively protect and monitor the marine environment across Europe, improving the knowledge on the consequences of stressors on marine biodiversity and on the functioning of coastal ecosystems, to sustainably manage the use of marine resources (Borja et al., 2010). Both initiatives combine sustainability and conservation for a responsible use of resources, through an ecosystem-based approach.

The structure and functioning of water ecosystems are essential for the application of these ecological-based regulations. The ecological status of ecosystems is assessed combining the physical, geographic and climatic factors of the natural habitat with the new altered conditions resulting from human impact

results (Bourlat et al., 2013). The assessment of ecological quality is based on established indices, using biological, hydro-morphological and physico-chemical elements. In biological elements, macroinvertebrate species are among the most used (Borja et al., 2000; Muxika et al., 2005). The monitor of benthic macroinvertebrate species is of primer importance because they are recognized as bioindicators (i.e. species that respond rapidly and are sensitive to environmental and human pressures; Johnston and Roberts, 2009), allowing to understand and monitor long-term variations in community structure and diversity (Teixeira et al., 2008; van Hoey et al., 2010). To undertake the identification of macrozoobenthic species and the implementation of such initiatives, integrative and rigorous monitoring strategies at different spatial and temporal scales are demanded.

1.3 Coastal macroinvertebrate biomonitoring

1.3.1 Challenges of monitoring macrozoobenthic communities

The complexity of the ocean makes sampling challenging, as well as the difficulty to develop long-term studies, limit our understanding in how marine ecosystems have been affected. Although in the last decades the knowledge of marine biodiversity has been increasing, marine species remain less studied than species from other ecosystems (e.g. terrestrial biodiversity). Few studies have targeted multiple sampling geographical regions at a large time scale (Dailianis et al., 2018).

The absence of effective sampling designs to capture the full taxonomic diversity could bring consequences on biodiversity losses (Hoegh-Guldberg and Bruno, 2010; Guidi et al., 2020). The complexity of sampling commonly results in limited biomonitoring programs in space and time, which may prevent the full taxonomic identification of a marine bulk sample (Borja et al., 2016; Costello et al., 2017). There are still difficulties in comparing monitoring studies since the inventories are often associated with low and unverifiable taxonomic precision. Additionally, coastal assessments are highly taxon oriented originating gaps for the same taxonomic groups, which are particularly poorly studied (Yeo et al., 2020).

Coastal hard-bottom substrates monitoring is commonly performed through scrapping the substrates (Wangensteen et al., 2018b; Shum et al., 2019; Turon et al., 2020), however a variety of other methods can also be used (Templado et al., 2010). This is challenging due to their three-dimensionally complex

and, most of the time, difficult access (Beisiegel et al., 2017). Furthermore, this is a destructive strategy, hard to standardize and hard to replicate within and among sampling campaigns, and normally implies loss of organisms (especially mobile fauna) (Obst et al., 2020). Artificial substrates have been used in ecological studies to assess biodiversity and to investigate their impact on species colonization, mostly due to their easier sampling process. The use of artificial substrates as sampling strategies have been used in different projects worldwide (e.g. David et al., 2019; Obst et al., 2020), and may provide a broad comparison between studies during next years. However, even using artificial substrates, within each sampled region, the number of sampling sites and replicates varied across studies (Leray and Knowlton, 2017).

Artificial substrates normally differ from natural substrates in composition, structure, complexity and orientation (Cacabelos et al., 2016). These established differences could influence the recruitment and settlement of species on the substrates, and the more complex an artificial substrate is, the higher the habitat available for colonization (García-Sanz et al., 2012). Roughness, composition and age of substrates were described as main influencers of community structure and diversity (e.g. Sedano et al., 2020). For example, plain surfaces of artificial substrates were correlated with lower colonization, particularly due to the lack of refuge leading to a negative impact on invertebrate larvae settlement (Koehl, 2007; Lee et al., 2004). Although some differences in the assemblages harbored by artificial substrates and natural communities have been reported (Sedano et al., 2019, 2020), other studies have demonstrated that artificial substrates promote the establishment of a community similar to the natural habitat (Cacabelos et al., 2016; Sedano et al., 2019).

The diversity of mobile communities is directly related and dependent on the primary colonization of new substrates by sessile organisms, creating perfect characteristics for species shelter. Across size fractions of macrozoobenthic species, the >2 mm mobile fraction is normally less diverse than smaller mobile fractions displaying differences in diversity between geographic regions (Leray and Knowlton, 2015). Moreover, sessile fauna had limited post-settlement dispersal abilities, making them more sensitive to differences in environmental factors at local scale. Additionally, after sieving and before DNA extraction, marine bulk samples are usually sorting into different size fractions (Leray and Knowlton, 2015; Ransome

et al., 2017), which could difficult the identification of small organisms being an extra time-consuming process.

1.3.2 Taxonomic identification methodologies

Macrobenthic species occurrence and abundance are routinely obtained through morphological-based identifications (Borja et al., 2009; Hering et al., 2010). However, identifying macrobenthic species can be challenging, which difficults the rapid assessment of human impacts and the implementation of efficient protection and restoration measures (Bourlat et al., 2013). For example, the lack of specific morphological characters and high degree of homoplasy difficult the visual distinction of poriferans (Bucklin et al., 2011).

The main morphological identifications drawbacks are associated to the high level of taxonomic expertise necessary to identify the species, resulting in a time-consuming, costly and unreliable approach (Yu et al., 2012). The presence of damaged specimens without diagnostic characters or early stages of development also difficult the assignment at species-level, resulting in low-throughput for biomonitoring (Lobo et al., 2017). Additionally, the use of inappropriated, incomplete or outdated taxonomic keys, combined with the continuous report of the presence of hidden or cryptic species (i.e. morphological similar, but genetically distinct) among macrozoobenthos result in underestimation of biodiversity when morphological-based approaches are used (Borges et al., 2016; Lobo et al., 2016). This is particularly important for marine macrozoobenthic communities, considering the high phylogenetic diversity of species, and may result in the impediment of a full diagnosis of a bulk community (Borja et al., 2016; Costello et al., 2017). The inability to perform large spatial and temporal studies, discarding a consistent qualitative taxonomic coverage, is also a shortfall of morphology-based assessments for biomonitoring (Bourlat et al., 2013).

The current morphology-related limitations may be circumvented by the application of new methodologies to assess and inventory species (Shokralla et al., 2012; Blackman et al., 2019). The emergence of DNA-based identification tools, such as DNA barcoding (i.e. the use of short sequences – the DNA barcodes - of a standardized zone of the eukaryotic genome as a molecular tag to generate DNA libraries for species identification) improve conventional approaches limitations (Hebert et al., 2003). This is particularly significant for DNA metabarcoding, which combine DNA barcoding with high-throughput sequencing

(Hajibabaei et al., 2011, 2012; Taberlet et al., 2012), allowing a rapid, reliable and scalable identification of multiple species from a single bulk sample. DNA metabarcoding brings several advantages over morphology-based identifications, emerging as a complementary methodology to study patterns of colonization by marine zoobenthic communities, eventually enabling higher throughput and efficiency in species detection and monitoring (Cowart et al., 2015; Elbrecht et al., 2017; Lobo et al., 2017). For example, using DNA metabarcoding higher taxonomic discrimination, i.e. at species-level, is commonly attained which greatly improve biomonitoring accuracy (Cristescu, 2014; Zinger et al., 2019). DNA-based methodologies are also being used to solve species and cryptic complexes (Govindarajan et al., 2006; Moura et al., 2008).

DNA metabarcoding allows the use of large scale and temporal strategies, obtaining fast results with a cost-efficiency higher than using morphology-based methodologies (Porter and Hajibabaei, 2018; Bush et al., 2019). Currently, some constraints in (meta)barcoding costs, due to insufficient funding for biodiversity studies, has been overcome driving the cost down and enabling a balanced effort to accurately identify species (Yang et al., 2018; Chang et al., 2020a). DNA-based approaches may provide good estimates of biodiversity, based on the identification of all biodiversity species from a macrozoobenthic sample improving the comprehensive assessment of marine ecosystems.

1.4 DNA-based approaches for macrozoobenthic biodiversity research and monitoring

1.4.1 DNA metabarcoding developed studies

DNA-based approaches have been widely implemented to assess and monitor macrozoobenthic diversity in coastal ecosystems. Recent studies show the applicability of molecular tools to characterize marine biodiversity and to facilitate environmental management protocols (Cahill et al., 2018; Cowart et al., 2018; Sawaya et al., 2019; Shum et al., 2019; Antich et al., 2020; Chang et al., 2020a, b; Huhn et al., 2020). In particular, complex hard-bottom communities can be monitored using DNA metabarcoding as a fast, robust, objective and affordable methodology for comprehensive characterization of marine macrobenthic diversity (Wangensteen et al., 2018b).

Different ecological questions have been also addressed through the implementation of DNA metabarcoding, enabling the monitoring of human-induced stressors. For example, to control fish farms (Keeley et al., 2018) or offshore oil-drilling activities (Lanzén et al., 2016), or to investigate the effects of oil spills on coastal environments (Xie et al., 2018). In a study of benthic macroinvertebrate communities from littoral stations, Aylagas et al. (2016) inferred a biotic index of ecological conditions and validates DNA metabarcoding as an environmental status assessment methodology. Also, DNA metabarcoding show promises for the identification of non-indigenous species, identifying new or earlier overlooked invasions and monitoring their effects on the habitats (Wangensteen et al., 2018a; Harper et al., 2020). The reveal of cryptic communities and small or rare taxa was been already demonstrated as an advantage of DNA metabarcoding over traditional morphology-based approaches (Al-Rshaidat et al., 2016; López-Escardó et al., 2018; Carvalho et al., 2019).

Other studies dedicated to investigate biogeographic patterns, seasonal and spatial environmental effects (Polinski et al., 2019; Pearman et al., 2020). Using coral reefs and artificial substrates, metabarcoding results showed higher sensitivity to identify differences between reef communities at small geographic scales, when compared to morphology (Perman et al., 2016). Indeed, most of the metabarcoding studies developed in marine ecosystems, particularly for hard-bottom substrates, have been performed using artificial substrates as sampling strategy (Plaisance et al., 2011; Al-Rshaidat et al., 2016; Leray and Knowlton, 2015, 2017; Ransome et al., 2017; Cahill et al., 2018; David et al., 2019). The applicability of DNA metabarcoding in biogeography, connectivity and dispersal studies and for conservation genetics was also demonstrated (Turon et al., 2020).

However, still fair to be implemented, through biodiversity assessments using DNA metabarcoding, the interconnection with ecological networks and machine learning is possible, improving our ability to detect community changes (Cordier et al., 2019; Fais et al., 2020).

1.4.2 Challenges of DNA metabarcoding approaches in coastal macrozoobenthic communities

Current weaknesses of DNA metabarcoding approaches includes lack of standardized protocols (sampling, DNA extraction and PCR amplification), incompleteness and inaccuracy of reference libraries and the absence of abundant data to calculate biotic indices (Cahill et al., 2018; McGee et al., 2019; Derycke et al., 2020). Furthermore, van der Loos and Nijland (2020) highlighted the lack of standardization in metabarcoding workflow, where the adopted methodologies often result on the specific and individual purposes of each study.

The preservation methods are essential to maintain the integrity and prevent degradation of the DNA from coastal bulk samples to capture the full taxonomic profile of the sample. For macrozoobenthic communities' preservation methods varied across studies. Although ethanol, combined with -20 ^{0}C, has been the preferred preservation method (Lobo et al., 2017; Aylagas et al., 2018), transport and shipping could be problematic due to safety restrictions and evaporation (Robinson et al., 2020). Alternative methods, particularly DMSO buffer and propylene grycol, have also been used as suitable alternatives without transport and shipping problems (Ransome et al., 2017; Robinson et al., 2020; Weigand et al., 2021). In a study using cryptic benthic communities where the authors tested different protocols, they found that DMSO results in samples with less degraded DNA and with higher quality (Ransome et al., 2017). However, a recent study found that the community profile detected in DMSO preserved samples was similar to the one derived from samples preserved in ethanol (Obst et al., 2020).

Before DNA extraction, samples are currently homogenized, avoiding the sorting step and reducing the processing time. However, this strategy prevents the benchmark between DNA metabarcoding based- with morphological based-identifications and favor the amplification of non-targeted taxa (e.g. non-metazoan reads) (Aylagas et al., 2018).

DNA extraction commercial kits have been used as the preferred methodology to extract all the DNA present in a marine bulk sample, allowing for a standard approach to recover high quality DNA, even in samples with high phylogenetic diversity (Duarte et al., *submitted*). Alternative methodologies, for example the direct extraction of the DNA from preservative ethanol combined with extraction kits, are more rapid and cost-effective in the characterization of macroinvertebrate communities (Martins et al., 2019; Zizka et al., 2019). Furthermore, non-destructive approaches allow for the preservation of all the specimens present in the

sample without any limit for the amount of material. However, in a recent study using marine macrobenthos, Derycke et al. (2020) detected less species in using preservative ethanol.

For DNA amplification of macroinvertebrate species, mitochondrial cytochrome oxidase subunit I (COI) gene have been the most used, especially due to the high representativeness of sequences in reference databases and by the ability to promote species-level identifications (Carew et al., 2013; Shokralla et al., 2015; Hollatz et al., 2017; Andújar et al., 2018; Weigand et al., 2019). Other marker loci broadly used is the nuclear small subunit 18S rRNA gene (18S rRNA), however species-level discrimination ability is normally lower than the obtained using COI (Tang et al., 2012; Lejzerowicz et al., 2015; Danovaro et al., 2016). The firstly designed primer-pairs for the entire COI-5P barcoding region (658 bp; Folmer et al., 1994), however, lack degenerate bases and due to the high taxonomic diversity of macrozoobenthic communities, mismatches between primers and sequences may occur frequently (Lobo et al., 2013). The choice of primers is crucial to attain highest levels of species detection since it will directly affect the species recovered from a bulk mixture, due to the different affinites with the sequences of the species composing that mixture. An enhanced primer-pair, targeting the 3' region of the COI-5P barcoding region of 313 bp (mlCOIintF/jgHCO2198; Leray et al., 2013), has been the most used in metabarcoding focused studies, with broad amplification for marine macrozoobenthos (Duarte et al., *submitted*). Other degenerate primers were designed for the same reverse region (e.g. LoboR1, Lobo et al., 2013) with high levels of success in zoobenthic species detection (Hollatz et al., 2017). Although different markers and primers can yield different diversity estimates, combining markers and/or primers could increase the ability to detect species. In a study using mock communities, Hollatz et al. (2017) was able to attain the highest species detection (85%), only when using at least three primer-pairs.

The accuracy of species-level assignments depends on comprehensive reference databases. Gaps in reference databases of barcode sequences for marine macrozoobenthic species have been reported for European waters (Weigand et al., 2019; Duarte et al., 2020), currently resulting in a high number of sequences unidentified, which may underestimate species diversity recovered through DNA metabarcoding. However taxonomic representativeness greatly varies among groups, and among marine macrozoobenthos the highest coverage of representative sequences in reference databases was found for

Annelida, Mollusca and Arthropoda (40-50%; Weigand et al., 2019), which are also the dominant taxonomic groups found in marine benthos.

Methodological optimizations still need to be concluded, to improve the sensitivity of DNA metabarcoding approaches to provide higher taxonomic discrimination and comparisons among studies and across a higher spatial and temporal scale for the bioassessment of coastal macrozoobenthic communities. However, comparison of DNA metabarcoding-based with morphology-based identifications may support the high-throughput and fast responses of DNA-based tools, adding comprehensive information of the communities for monitoring coastal ecosystems.

1.5 Aims and structure of the thesis

DNA-based methodologies have been increasingly applied to assess the species composition in a variety of communities and ecosystems, revolutionizing our ability to observe patterns and trends in biodiversity. DNA metabarcoding provides the opportunity to augment throughput and improve the ability to accurately and cost-effectively monitor coastal macrozoobenthic communities, particularly when coupled with standardized sampling strategies (i.e. artificial substrates). Despite several trial studies that have shown this potential, some important challenges remain. Among others, the risk that systematic shortfalls are introduced in the assessments can lead to missing the detection of important fractions of the communities' diversity. Indeed, it is still necessary to optimize procedures to assure that technical biases are minimized, and to increase standardization and comparability among studies. Therefore, the main goal of the present thesis is to develop and optimize a DNA metabarcoding-based system, coupled with artificial substrates, for monitoring marine macrozoobenthic communities. This system is meant to provide a high-throughput approach that captures a near-complete profile of the species diversity of the community. We envisage that such a system can be eventually implemented under routine monitoring for European WFD, MSFD, and other purposes, like broad-scale assessment of the impact of global change in macrozoobenthic communities, species ranges and interactions. Aiming to generate baseline data for future reference, we also aimed to investigate overtime patterns of macrobenthic colonization in artificial substrates deployed in selected locations of the Northwest Iberian Atlantic coast. In order to accomplish these goals the following studies were performed:

- Gap-analyses, assembly and annotation of a comprehensive reference library for the marine macrozoobenthic species of the Atlantic Iberia;
- Optimization of DNA metabarcoding protocols, namely DNA extraction and amplification protocols, investigating the ability of different mitochondrial and nuclear target loci and primer-pairs to amplify, and therefore enable detection, of the full spectrum of species diversity of marine macrozoobenthos;
- Comparison of the performance of morphology and DNA metabarcoding-based methodologies for species detection and taxonomic profiling of the macrozoobenthos;
- Testing the effect of artificial substrates materials (slate, PVC and granite) and shapes (ARMS, ASMS) on the composition and patterns of colonization of marine macrozoobenthic communities;
- Investigation of the temporal and regional patterns of colonization of macrozoobenthic species in artificial substrates deployed at selected locations in Northwest Iberian coast.

This thesis is divided in six chapters, four of which (Chapters 2 to 5) consist on the experimental studies performed in the scope of this thesis and organized in individual sections (Fig. 1.1). All studies were performed under the scope of the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF), by the project "The NextSea: Next generation monitoring of coastal ecosystems in a scenario of global change" (NORTE-01-0145-FEDER-000032). Financial support for the thesis author was attained through an FCT fellowship (PD/BD/127994/2016), hosted by the University of Minho (Portugal) and the University of Vigo (Spain), in the scope of the PhD programme in Marine Science Technology and Management (Do*Mar).

Chapter 1 corresponds to the general introduction, including a short literature review on the thesis underlying topics, and the thesis goals and structure. In Chapter 2 a comprehensive reference library of DNA barcodes was assembled for marine macroinvertebrate species from Atlantic Iberia, together with the assessment of gaps in the representativeness of species sequences as well as the examination of data ambiguities. Chapter 3 reports on the optimization of the protocol for DNA metabarcoding assessment of coastal macrozoobenthic communities, where the efficiency of two marker loci and different primer-pairs

on species-level taxonomic assignment is evaluated, and the representativeness of the species detected in reference libraries is also investigated. In Chapters 4 and 5 we investigated the influence of artificial substrates on the taxonomic composition and structure of marine macrozoobenthic communities. In Chapter 4 we tested artificial substrates made of different materials (slate, PVC and granite) that were deployed at one location in the Northwest Iberian coast (Toralla Island, Ría de Vigo). In this chapter, morphology and DNA metabarcoding approaches were compared first, and combined later, to assess the taxonomic composition of macrozoobenthos at different time intervals since deployment of the substrates. In Chapter 5 the influence of the shape of artificial substrates on species colonization is investigated using two different types of substrates: ARMS and ASMS. The substrates were deployed at two locations in NW Iberian coast (near the cities of Vigo and Ferrol), and DNA metabarcoding applied to assess the taxonomic composition of macrozoobenthos, and to study the patterns of temporal variation at regional level. Chapter 6 consists in the global appraisal of the thesis findings, with the concluding remarks and future perspectives. The three annexes integrated in this thesis consist on the supplementary information and additional figures and tables from chapters 3 to 5. Each of the thesis experimental chapters correspond to one scientific article, which has been published or is in preparation for publication in due course:

Chapter 2

Leite B.R., Vieira P.E., Teixeira M.A.L., Lobo-Arteaga J., Hollatz C., Borges L.M.S., Duarte S., Troncoso J.S., Costa F.O. (2020). Gap-analysis and annotated reference library for supporting macroinvertebrate metabarcoding in Atlantic Iberia. Regional Studies in Marine Science. 36, 101307. https://doi.org/10.1016/j.rsma.2020.101307.

Chapter 3

Leite B.R., Vieira P.E., Troncoso J.S., Costa F.O. (*in prep*). Comparing species detection success between molecular markers and primers in DNA metabarcoding of coastal macroinvertebrate communities.

Chapter 4

Leite B.R., Troncoso J.S., Costa F.O. (*in prep*). Coastal macrozoobenthos monitoring: impact of artificial substrate material and comparison of morphology vs DNA metabarcoding-based taxonomic profiling.

Chapter 5

Leite B.R., Troncoso J.S., Costa F.O. (*in prep*). ARMS embracing ASMS to complement DNA metabarcoding-based monitoring of macrozoobenthos in NW Iberian coast.

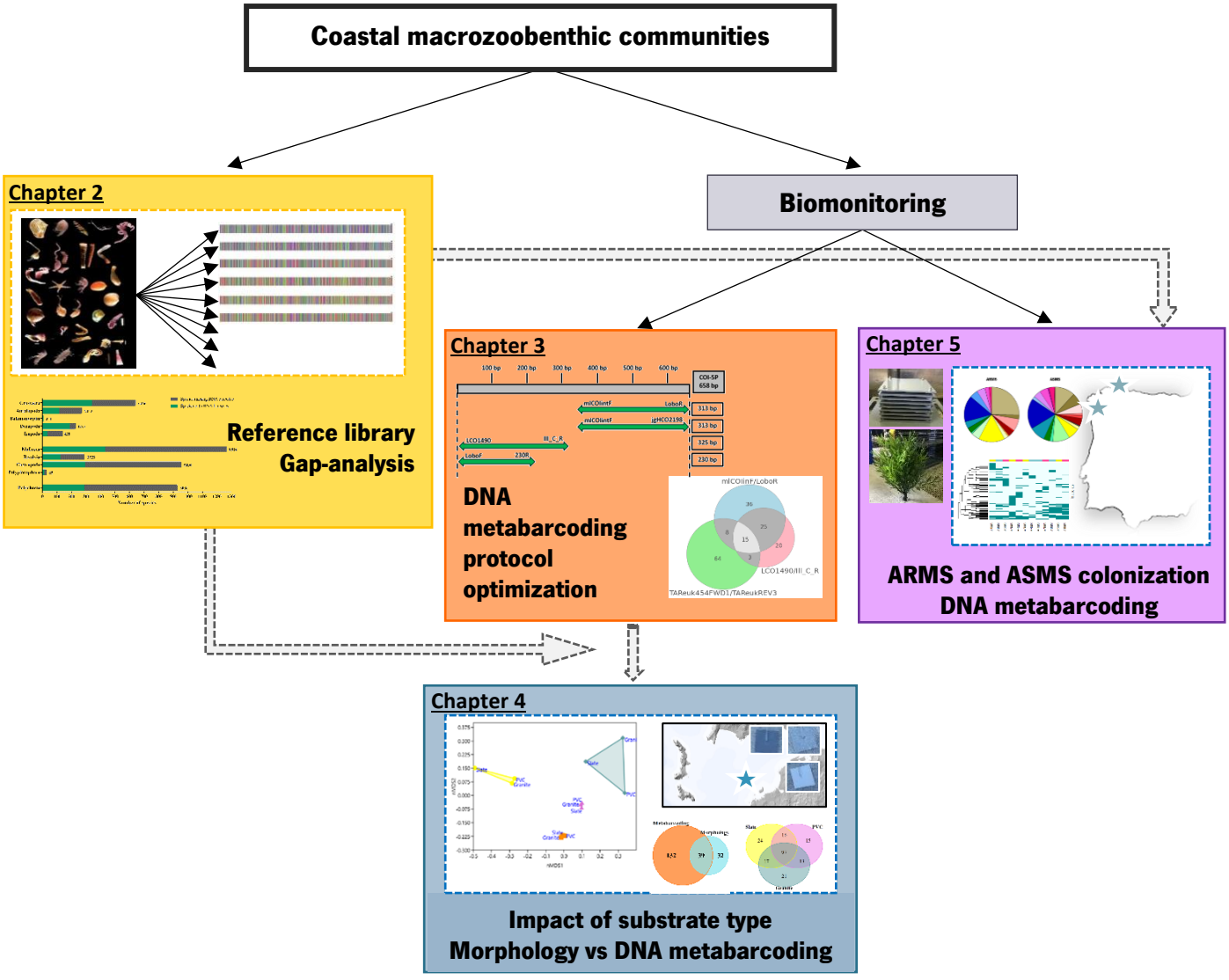


Figure 1.1. Representative scheme of the structure of the thesis, with the connections between chapters.

References

Albano P.G., Sabelli B., Bouchet P. (2011). The challenge of small and rare species in marine biodiversity surveys: microgastropod diversity in a complex tropical coastal environment. Biodiversity and Conservation. 20(13), 3223–3237. https://doi.org/10.1007/s10531-011-0117-x.

Al-Rshaidat M.M.D., Snider A., Rosebraugh S., Devine A.M., Devine T.D., Plaisance L., Knowlton N., 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. https://doi.org/10.1139/gen-2015-0208.

Andújar C., Arribas P., Yu D.W., Vogler A.P., Emerson B.C. (2018). Why the COI barcode should be the community DNA metabarcode for the metazoa. Molecular Ecology. 27, 3968–3975. https://doi.org/10.1111/mec.14844.

Antich A., Palacín C., Cebrian E., Golo R., Wangensteen O.S., Turon X. (2020). Marine biomonitoring with eDNA: Can metabarcoding of water samples cut it as a tool for surveying benthic communities? Molecular Ecology. https://doi.org/10.1111/mec.15641.

Appeltans W., Ahyong S.T., Anderson G., Angel M.V., Artois T., Bailly N., Bamber R., Barber A., Bartsch I., Berta A., BlazewiczPaszkowycz M., Bock P., Boxshall G., Boyko C.B., Brandão S.N., Bray R.A., Bruce N.L., Cairns S.D., Chan T., Cheng L., Collins A.G., Cribb T., Curini-Galletti M., Dahdouh-Guebas F., Davie P.J.F., Dawson M.N., de Cleck O., Decock W., de Grave S., de Voogd N.J., Domning D.P., Emig C.C., Erséus C., Eschmeyer W., Fauchald K., Fautin D.G., Feist S.W., Fransen C.H.J.M., Furuya H., Garcia-Alvarez O., Gerken S., Gibson D., Gittenberger A., Gofas S., Gómez-Daglio L., Gordon D.P., Guiry M.D., Hernandez F., Hoeksema B.W., Hopcroft R.R., Jaume D., Kirk P., Kowdam N., Koenemann S., Kolb J.B., Kristensen R.M., Kroh A., Lambert G., Lazarus D.B., Lemaítre R., Longshaw M., Lowry J., Macpherson E., Madin L.P., Mah C., Mapstone G., McLaughlin P.A., Mees J., Meland K., Messing C.G., Mills C.E., Molodtsova T.N., Mooi R., Neuhaus B., Ng P.K.L., Nielsen C., Norenburg J., Opresko D.M., Osawa M., Paulay G., Perrin W., Pilger J.F., Poore G.C.B., Pugh P., Read G.B., Reimer J.D., Rius M., Rocha R.M., Saiz-Salinas J.I., Scarabino V., Schierwater B., Schmidt-Rhaesa A., Schnabel K.E., Schotte M., Schubert P., Schwabe E., Segers H., Self-Sullivan C., Shenkar N., Siegel V., Sterrer W., Stöhr S., Swalla B., Tasker M.L., Thuesen E.V., Timm T., Todaro M.A., Turon X., Tyler S., Uetz P., van der Land J., Vanhoorne B., van Ofwegen L.P., van Soest R.W.M, Vanaverbeke J., Walker-Smith G., Walter T.C., Watten A., Williams G.C., Wilson S.P., Costello M.J. (2012). The magnitude of global marine species diversity. Current Biology. 22(23), 2189–2202. https://doi.org/10.1016/j.cub.2012.09.036.

Aylagas E., Borja A., Irigoien X., Rodriguez-Ezpeleta N. (2016). Benchmarking DNA metabarcoding for biodiversity-based monitoring and assessment. Frontiers in Marine Science. 3, 96. https://doi.org/10.3389/fmars.2016.00096.

Aylagas E., Borja Á., Muxika I., Rodríguez-Ezpeleta N. (2018). Adapting metabarcoding-based benthic biomonitoring into routine marine ecological status assessment networks. Ecological Indicators. 95, 194–202. https://doi.org/10.1016/j.ecolind.2018.07.044.

Barbier E.B. (2012). Progress and challenges in valuing coastal and marine ecosystem services. Review of Environmental Economics and Policy. 6, 1–19. https://doi.org/10.1093/reep/rer017.

Barbier E.B. (2017). Marine ecosystem services. Current Biology. 27(11), R507-R510. https://doi.org/10.1016/j.cub.2017.03.020.

Beisiegel K., Darr A., Gogina M., Zettler M.L. (2017). Benefits and shortcomings of non-destructive benthic imagery for monitoring hard-bottom habitats. Marine Pollution Bulletin. 121, 5–15. https://doi.org/10.1016/j.marpolbul.2017.04.009.

Blackman R.C., Mächler E., Altermatt F., Arnold A., Beja P., Boets P., Egeter B., Elbrecht V., Filipe A.F., Jones J.I., Macher J., Markus M., Martins F.M.S., Cesc M., Kristian M., Pawlowski J., Schmidt P.L.Y., Zizka V.M.A., Leese F., Price B.W., Kristy D. (2019). Advancing the use of molecular methods for routine freshwater macroinvertebrate biomonitoring – the need for calibration experiments. Metabarcoding and Metagenomics. 3, e34735. https://doi.org/10.3897/mbmg.3.34735.

Borges L.M.S., Hollatz C., Lobo J., Cunha A.M., Vilela A.P., Calado G., Coelho R., Costa A.C., Ferreira M.S.G., Costa M.H., Costa F.O. (2016). With a little help from DNA barcoding: investigating the diversity of Gastropoda from the Portuguese coast. Scientific Reports. 6, 20226. https://doi.org/10.1038/srep20226.

Borja A., Elliott M., Andersen J.H., Berg T., Carstensen J., Halpern B.S., Heiskanen A.S., Korpinen S., Lowndes J.S.S., Martin G., Rodríguez-Ezpeleta N. (2016). Overview of integrative assessment of marine systems: the ecosystem approach in practice. Frontiers in Marine Science. 3, 20. https://doi.org/10.3389/fmars.2016.00020.

Borja Á., Elliott M., Carstensen J., Heiskanen A.S., van de Bund W. (2010). Marine management – towards an integrated implementation of the European Marine Strategy Framework and the Water Framework Directives. Marine Pollution Bulletin. 60(12), 2175-2186. https://doi.org/10.1016/j.marpolbul.2010.09.026.

Borja A., Franco J., Pérez V. (2000). A marine biotic index to establish the ecological quality of soft-bottom benthos within European estuarine and coastal environments. Marine Pollution Bulletin. 40, 1100–1114. https://doi.org/10.1016/S0025-326X(00)00061-8.

Borja A., Miles A., Occhipinti-Ambrogi A., Berg T. (2009). Current status of macroinvertebrate methods used for assessing the quality of European marine waters: implementing the Water Framework Directive. Hydrobiologia. 633, 181–196. https://doi.org/10.1007/s10750-009-9881-y.

Bouchet P. (2006). The magnitude of marine biodiversity. In Duarte C.M. The exploration of marine biodiversity, Scientific and technological challenges. Fundación BBVA.

Bourlat S.J., Borja A., Gilbert J., Taylor M.I., Davies N., Weisberg S.B., Griffith J.F., Lettieri T., Field D., Benzie J., Glöckner F.O., Rodríguez-Ezpeleta N., Faith D.P., Obst M. (2013). Genomics in marine monitoring: new opportunities for assessing marine health status. Marine Pollution Bulletin. 74(1), 19-31. http://dx.doi.org/10.1016/j.marpolbul.2013.05.042.

Bucklin A., Steinke D., Blanco-Bercial L. (2011). DNA barcoding of marine metazoa. Annual review of Marine Science. 3, 471-508. http://doi.org/10.1146/annurev-marine-120308-080950.

Bush A., Compson Z.G., Monk W.A., Porter T.M., Steeves R., Emilson E., Gagne N., Hajibabaei M., Roy M., Baird D.J. (2019). Studying ecosystems with DNA metabarcoding: lessons from biomonitoring of aquatic. Frontiers in Ecology and Evolution. 7, 434. https://doi.org/10.3389/fevo.2019.00434.

Cacabelos E., Martins G.M., Thompson R., Prestes A.C.L., Azevedo J.M.N., Neto A.I. (2016). Material type and roughness influence structure of inter-tidal communities on coastal defenses. Marine Ecology. 37, 801–812. https://doi.org/10.1111/maec.12354.

Cahill A.E., Pearman J.K., Borja A., Carugati L., Carvalho S., Danovaro R., Dashfield S., David R., Féral J., Olenin S., Šiaulys A., Somerfield P.J., Trayanova A., Uyarra M.C., Chenuil A. (2018). A comparative analysis of metabarcoding and morphology-based identification of benthic communities across different regional seas. Ecology and Evolution. 8, 8908–8920. https://doi.org/10.1002/ece3.4283.

Carew M.E., Pettigrove V.J., Metzeling L., Hoffmann A.A. (2013). Environmental monitoring using next generation sequencing: rapid identification of macroinvertebrate bioindicator species. Frontiers in Zoology. 10, 45. https://doi.org/10.1186/1742-9994-10-45.

Carvalho S., Aylagas E., Villalobos R., Kattan Y., Berumen M., Pearman J.K. (2019). Beyond the visual: using metabarcoding to characterize the hidden reef cryptobiome. Proceedings of the Royal Society B Biological Sciences. 286, 20182697. https://doi.org/10.1098/rspb.2018.2697.

Chang J.J.M., Ip Y.C.A., Bauman A.G., Huang D. (2020a). MinION-in-ARMS: Nanopore sequencing to expedite barcoding of specimen-rich macrofaunal samples from Autonomous Reef Monitoring Structures. Frontiers in Marine Science. 7, 448. https://doi.org/10.3389/fmars.2020.00448.

Chang J.J.M., Ip Y.C.A., Ng C.S.L., Huang D. (2020b). Takeaways from mobile DNA barcoding with bentoLab and MinION. Genes. 11(10), 1121. https://doi.org/10.3390/genes11101121.

Chapman M.G. (2002). Early colonization of shallow subtidal boulders in two habitats. Journal of Experimental Marine Biology and Ecology. 275, 95–116. https://doi.org/10.1016/S0022-0981(02)00134-X.

Chapman M.G. (2003). Paucity of mobile species on constructed seawalls: effects of urbanization on biodiversity. Marine Ecology Progress Series. 264, 21-29. https://doi.org/10.3354/meps264021.

Chapman M.G., Underwood A.J., Browne M.A. (2018). An assessment of the current usage of ecological engineering and reconciliation ecology in managing alterations to habitats in urban estuaries. Ecological Engineering. 120, 560-573. https://doi.org/10.1016/j.ecoleng.2017.06.050.

Clements J.C., Darrow E.S. (2018). Eating in an acidifying ocean: a quantitative review of elevated CO_2 effects on the feeding rates of calcifying marine invertebrates. Hydrobiologia. 820(1), 1–21. https://doi.org/10.1007/s10750-018-3665-1.

Cordier T., Lanzén A., Apothéloz-Perret-Gentil L., Stoeck T., Pawlowski J. (2019). Embracing environmental genomics and machine learning for routine biomonitoring. Trends in Microbiology. 27, 387–397. https://doi.org/10.1016/j.tim.2018.10.012.

Costello M.J., Basher Z., McLeod L., Asaad I., Claus S., Vandepitte L., Yasuhara M., Gislason H. Edwards M., Appeltans W., Enevoldsen H., Edgar G.J., Miloslavich P., de Monte S., Sousa-Pinto I., Obura D., Bates A.E. (2017). Methods for the study of marine biodiversity. In Walters M., Scholes R.J. The GEO Handbook on biodiversity observation networks. Springer International Publishing. ISBN 978-3-319-27288-7.

Cowart D.A., Murphy K.R., Cheng C.H.C. (2018). Metagenomic sequencing of environmental DNA reveals marine faunal assemblages from the West Antarctic Peninsula. Marine Genomics. 37, 148-160. https://doi.org/10.1016/j.margen.2017.11.003.

Cowart D.A., Pinheiro M., Mouchel O., Maguer M., Grall J., Miné J., Arnaud-Haond S. (2015). Metabarcoding is powerful yet still blind: a comparative analysis of morphological and molecular surveys of seagrass communities. PLoS One. 10, e0117562. https://doi.org/10.1371/journal.pone.0117562.

Cristescu M.E. (2014). From barcoding single individuals to metabarcoding biological communities: towards an integrative approach to the study of global biodiversity. Trends in Ecology and Evolution. 29, 566–571. https://doi.org/10.1016/j.tree.2014.08.001.

Culhane F.E., Frid C.L., Gelabert E.R., White L., Robinson L.A. (2018). Linking marine ecosystems with the services they supply: what are the relevant service providing units? Ecological Applications. 28(7), 1740-1751. https://doi.org/10.1002/eap.1779.

Dailianis T., Smith C.J., Papadopoulou N., Gerovasileiou V., Sevastou K., Bekkby T., Bilan M., Billet D., Boström C., Carreiro-Silva M., Danovaro R., Fraschetti S., Gagnon K., Gambi C., Grehan A., Kipson S., Kotta J., McOwen C.J., Morato T., Ojaveer H, Pham C.K., Scrimgeour R. (2018). Human activities and resultant pressures on key European marine habitats: an analysis of mapped resources. Marine Policy. 98, 1–10. https://doi.org/10.1016/j.marpol.2018.08.038.

Danovaro R., Carugati L., Berzano M., Cahill A.E., Carvalho S., Chenuil A., Corinaldesi C., Cristina S., David R., Dell'anno A., Dzhembekova N., Garcés E., Gasol J., Goela P., Feral J.P., Ferrera I., Forster R., Kurekin A.A., Rastelli E., Marinova V., Miller P.I., Moncheva S., Newton, A., Pearman J., Pitois S., Renè A., Rodriguez-Ezpeleta N., Saggiomo V., Simis S.G.H., Stefanova K., Wilson C., Lo Martire M., Greco S., Cochrane S., Borja A. (2016). Implementing and innovating marine monitoring approaches for assessing marine environmental status. Frontiers in Marine Science. 3, 213. https://doi.org/10.3389/fmars.2016.00213.

David R., Uyarra M.C., Carvalho S., Anlauf H., Borja A., Cahill A.E., Carugati L., Danovaro R., De Jode A., Féral J., Guillemain D., Lo Martire M., D'Avray L.T.V., Pearman J.K., 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. https://doi.org/10.1016/j.marpolbul.2019.02.066.

Derycke S., Maes S., Hillewaert H., Ampe B., Haegeman A., Hostens K., De A. (2020). Optimisation of metabarcoding for monitoring marine macrobenthos: primer choice and morphological traits determine species detection in bulk DNA and eDNA from the ethanol preservative. Authorea. https://doi.org/10.22541/au.159665093.39993653.

Duarte S., Leite B.R., Feio M.J., Costa F.O., Filipe A.F. (submitted). Integration of DNA-based approaches in aquatic ecological assessment using benthic macroinvertebrates. Water.

Duarte S., Vieira P.E., Costa F.O. (2020). Assessment of species gaps in DNA barcode libraries of non-indigenous species (NIS) occurring in European coastal regions. Metabarcoding and Metagenomics. 4, e55162. https://doi.org/10.3897/mbmg.4.55162.

Elbrecht V., Vamos E.E., Meissner K., Aroviita J., Leese F. (2017). Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. Methods Ecology and Evolution. 8, 1265–1275. https://doi.org/10.1111/2041-210X.12789.

European Commission, 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 Establishing a Framework for Community Action in the Field of Water Policy. https://eur-lex.europa.eu/eli/dir/2000/60.

European Commission, 2008. Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 Establishing a Framework for Community Action in the Field of Marine Environmental Policy (Marine Strategy Framework Directive). https://eur-lex.europa.eu/eli/dir/2008/56.

Fais M., Bellisario B., Duarte S., Vieira P.E., Sousa R., Canchaya C., Costa F.O. (2020). Meiofauna metabarcoding in Lima estuary (Portugal) suggests high taxon replacement within a background of network stability. Regional Studies in Marine Science. 38, 101341. https://doi.org/10.1016/j.rsma.2020.101341.

Firth L.B., Knights A.M., Bridger D., Evans A., Mieskowska N., Moore P.J., O'Connor N.E., Sheehan E., Thompson R.C., Hawkins S.J. (2016). Ocean sprawl: challenges and opportunities for biodiversity management in a changing world. Oceanography and Marine Biology: An Annual Review. 54, 189-262. https://doi.org/10.1201/9781315368597.

Folmer O., Black M., Hoeh W., Lutz R., Vrijenhoek R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology. 3(5), 294–9.

Friedrichs M., Graf G., Springer B. (2000). Skimming flow induced over a simulated polychaete tube lawn at low population densities. Marine Ecology Progress Series. 192, 219–228. https://doi.org/10.3354/meps192219.

Gage J.D. (2001). Macrobenthos. Encyclopedia of Ocean Sciences. 3, 1505-1515. https://doi.org/10.1016/B978-012374473-9.00211-3.

Gallego R., Jacobs-Palmer E., Cribari K., Kelly R.P. (2020). Environmental DNA metabarcoding reveals winners and losers of global change in coastal waters. Proceedings of the Royal Society B. 287(1940), 2020-2424. https://doi.org/10.1098/rspb.2020.2424.

García-Sanz S., Tuya F., Navarro P.G., Angulo-Preckler C., Haroun R.J. (2012). Post larval, short-term, colonization patterns: The effect of substratum complexity across subtidal, adjacent, habitats. Estuarine, Coastal and Shelf Science. 112, 183-191. https://doi.org/10.1016/j.ecss.2012.07.014.

Govindarajan A.F., Boero F., Halanych K.M. (2006). Phylogenetic analysis with multiple markers indicates repeated loss of the adult medusa stage in Campanulariidae (Hydrozoa, Cnidaria). Molecular Phylogenetics and Evolution. 38(3), 820–34. https://doi.org/10.1016/j.ympev.2005.11.012.

Grabowski J.H., Brumbaugh R.D., Conrad R.F., Keeler A.G., Opaluch J.J., Peterson C.H., Piehler M.F., Powers S.P., Smyth, A. R. (2012). Economic valuation of ecosystem services provided by oyster reefs. Bioscience. 62(10), 900-909. https://doi.org/10.1525/bio.2012.62.10.10.

Guerra-García J.M., Baeza-Rojano E., Cabezas M.P., García-Gómez J.C. (2011). Vertical distribution and seasonality of peracarid crustaceans associated with intertidal macroalgae. Journal of Sea Research. 65(2), 256-264. https://doi.org/10.1016/j.seares.2010.12.001.

Guidi L., Guerra A.F, Canchaya C., Curry E., Foglini F., Irisson J.O., Malde K., Marshall C.T., Obst M., Ribeiro R.P., Tjiputra J., Bakker D.C.E. (2020). Big data in marine science. In Alexander B., Heymans J.J., Piniella A.M., Kellett P., Coopman J. Future Science Brief 6. European Marine Board. https://dx.doi.org/10.5281/zenodo.3755793.

Hajibabaei M., Shokralla S., Zhou X., Singer G.A.C., Baird D.J. (2011). Environmental barcoding: a Next-Generation Sequencing approach for biomonitoring applications using river benthos. PLoS One. 6, e17497. https://doi.org/10.1371/journal.pone.0017497.

Hajibabaei M., Spall J.L., Shokralla S., van Konynenburg S. (2012). Assessing biodiversity of a freshwater benthic macroinvertebrate community through non-destructive environmental barcoding of DNA from preservative ethanol. BioMed Central Ecology. 12, 28. https://doi.org/10.1186/1472-6785-12-28.

Harper L.R., Lawson Handley L., Sayer C.D., Read D.S., Benucci M., Blackman R.C., Hill M.J., Hänfling B. (2020). Assessing the impact of the threatened crucian carp (*Carassius carassius*) on pond invertebrate diversity: A comparison of conventional and molecular tools. Molecular Ecology. https://doi.org/10.1111/mec.15670.

Hebert P.D.N., Cywinska A., Ball S.L., deWaard J.R. (2003). Biological identifications through DNA barcodes. Proceedings of the Royal Society of London. Series B: Biological Sciences. 270(1512), 313-321. https://doi.org/10.1098/rspb.2002.2218.

Hereu B., Zabala M., Linares C., Sala E. (2005). The effects of predator abundance and habitat structural complexity on survival of juvenile sea urchins. Marine Biology. 146, 293–299. https://doi.org/10.1007/s00227-004-1439-y.

Hering D., Borja A., Carstensen J., Carvalho L., Elliott M., Feld C.K., Heiskanen A.S., Johnson R.K., Moe J., Pont D., Solheim A.L., van de Bund W. (2010). The European Water Framework Directive at the age of 10: a critical review of the achievements with recommendations for the future. Science of the Total Environment. 408, 4007–4019. https://doi.org/10.1016/j.scitotenv.2010.05.031.

Hoegh-Guldberg O., Bruno J.F. (2010). The impact of climate change on the world's marine ecosystems. Science. 328(5985), 1523-1528. https://doi.org/10.1126/science.1189930.

Hollatz C., Leite B.R., Lobo J., Froufe H., Egas C., Costa F.O. (2017). Priming of a DNA metabarcoding approach for species identification and inventory in marine macrobenthic communities. Genome. 60, 260–271. https://doi.org/10.1139/gen-2015-0220.

Huhn M., Madduppa H.H., Khair M., Sabrian A., Irawati Y., Anggraini N.P., Wilkinson S.P., Simpson T., Iwasaki K., Setiamarga D.H.E., Dias P.J. (2020). Keeping up with introduced marine species at a remote biodiversity hotspot: awareness, training and collaboration across different sectors is key. Biological Invasions. 22(2), 749-771. https://doi.org/10.1007/s10530-019-02126-2.

Jackson J.B., Kirby M.X., Berger W.H., Bjorndal K.A., Botsford L. W., Bourque B.J., Bradbury R.H., Cooke R., Erlandson J., Estes J.A., Hughes T.P., Kidwell S., Lange C.B., Lenihan H.S., Pandolfi J.M., Peterson C.H., Steneck R.S., Tegner M.J., Warner R.R. (2001). Historical overfishing and the recent collapse of coastal ecosystems. Science. 293(5530), 629-637. https://doi.org/10.1126/science.1059199.

Jenkins S.R., Marshall D., Fraschetti S. (2009). Settlement and recruitment. In Wahl M. Marine hard bottom communities. Ecological Studies (Analysis and Synthesis). 206. Springer. https://doi.org/10.1007/b76710_12.

Johnston E.L., Roberts D.A. (2009). Contaminants reduce the richness and evenness of marine communities: a review and meta-analysis. Environmental Pollution. 157(6), 1745–1752. https://doi.org/10.1016/j.envpol.2009.02.017.

Keeley N., Wood S.A., Pochon X. (2018). Development and preliminary validation of a multi-trophic metabarcoding biotic index for monitoring benthic organic enrichment. Ecological Indicators. 85, 1044–1057. https://doi.org/10.1016/j.ecolind.2017.11.014.

Koehl M.R.A. (2007). Mini review: hydrodynamics of larval settlement into fouling communities. Biofouling. 23, 357-368. https://doi.org/10.1080/08927010701492250.

Lalli C., Parsons T.R. (1997). Biological oceanography: an introduction. Elsevier.

Lanzén A., Lekang K., Jonassen I., Thompson E.M., Troedsson C. (2016). High-throughput metabarcoding of eukaryotic diversity for environmental monitoring of offshore oil-drilling activities. Molecular Ecology. 25, 4392–4406. https://doi.org/10.1111/mec.13761.

Lau J.D., Hicks C.C., Gurney G.G., Cinner J.E. (2019). What matters to whom and why? Understanding the importance of coastal ecosystem services in developing coastal communities. Ecosystem Services. 35, 219-230. https://doi.org/10.1016/j.ecoser.2018.12.012.

Lee J.T., Widdows J., Jones M.B., Coleman R.A. (2004). Settlement of megalopae and early juveniles of the velvet swimming crab *Necora puber* (Decapoda: Portunidae) in flow conditions. Marine Ecology Progress Series. 272, 191–202. https://doi.org/10.3354/meps272191.

Lejzerowicz F., Esling P., Pillet L., Wilding T.A. (2015). High-throughput sequencing and morphology perform equally well for benthic monitoring of marine ecosystems. Scientific Reports. 5, 13932. https://doi.org/10.1038/srep13932.

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. 112(7), 2076–2081. https://doi.org/10.1073/pnas.1424997112.

Leray M., Knowlton N. (2017). Random sampling causes the low reproducibility of rare eukaryotic OTUs in Illumina COI metabarcoding. PeerJ. 5, e3006. https://doi.org/10.7717/peerj.3006.

Leray M., Yang J.Y., Meyer C.P., Mills S.C., Agudelo N., Ranwez V., Boehm J.T., 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. https://doi.org/10.1186/1742-9994-10-34.

Liquete C., Piroddi C., Drakou E.G., Gurney L., Katsanevakis S., Charef A., Ego B. (2013). Current status and future prospects for the assessment of marine and coastal ecosystem services: a systematic review. PloS One. 8(7), e67737. https://doi.org/10.1371/journal.pone.0067737.

Lobo J., Costa P.M., Teixeira M.A.L., Ferreira M.S.G, Costa M.H., Costa F.O. (2013). Enhanced primers for amplification of DNA barcodes from a broad range of marine metazoans. BMC Ecology. 13, 34. https://doi.org/10.1186/1472-6785-13-34.

Lobo J., Shokralla S., Costa M.H., Hajibabaei M., Costa F.O. (2017). DNA metabarcoding for high-throughput monitoring of estuarine macrobenthic communities. Scientific Reports. 7, 15618. https://doi.org/10.1038/s41598-017-15823-6.

Lobo J., Teixeira M.A.L., Borges L.M., Ferreira M.S.G., Hollatz C., Gomes P.T., Sousa R., Ravara A., Costa M.H., Costa F.O. (2016). Starting a DNA barcode reference library for shallow water polychaetes from the southern European Atlantic coast. Molecular Ecology Resources. 16 (1), 298–313. http://dx.doi.org/10.1111/ 1755-0998.12441.

López-Escardó D., Paps J., de Vargas C., Massana R., Ruiz-Trillo I., del Campo J. (2018). Metabarcoding analysis on European coastal samples reveals new molecular metazoan diversity. Scientific Reports. 8, 910. https://doi.org/10.1038/s41598-018-27509-8.

Lundberg J., Moberg F. (2003). Mobile link organisms and ecosystem functioning: implications for ecosystem resilience and management. Ecosystems. 6, 87–98. https://doi.org/10.1007/s10021-002-0150-4.

Martins F.M.S., Galhardo M., Filipe A.F., Teixeira A., Pinheiro P., Paupério J., Alves P.C., Beja P. (2019). Have the cake and eat it: Optimizing nondestructive DNA metabarcoding of macroinvertebrate samples for freshwater biomonitoring. Molecular Ecology Resources. 19, 863–876. https://doi.org/10.1111/1755-0998.13012.

McGee K.M., Robinson C.V. Hajibabaei M. (2019). Gaps in DNA-based biomonitoring across the globe. Frontiers in Ecology and Evolution. 7, 337. https://doi.org/10.3389/fevo.2019.00337.

McLachlan A., Defeo O. (2017). The ecology of sandy shores. Elsevier. https://doi.org/10.1016/B978-0-12-809467-9.00007-2.

Molnar J.L., Gamboa R.L., Revenga C., Spalding M.D. (2008). Assessing the global threat of invasive species to marine biodiversity. Frontiers in Ecology and the Environment. 6, 485–492. https://doi.org/10.1890/070064.

Mora C., Tittensor D.P., Adl S., Simpson A.G.B., Worm B. (2011). How many species are there on Earth and in the Ocean? PloS Biology. 9(8), e1001127. https://doi.org/10.1371/journal.pbio.1001127.

Moreira J., Lourido A., Troncoso J.S. (2008). Diversity and distribution of peracarid crustaceans in shallow subtidal soft bottoms at the Ensenada de Baiona (Galicia, NW Spain). Crustaceana. 81, 1069–1089. https://doi.org/10.1163/156854008X360815.

Moura C.J., Harris D.J., Cunha M.R., Rogers A.D. (2008). DNA barcoding reveals cryptic diversity in marine hydroids (Cnidaria, Hydrozoa) from coastal and deep-sea environments. Zoologica Scripta. 31, 93–108. https://doi.org/10.1111/j.1463-6409.2007.00312.x.

Muxika I., Borja A., Bonne W. (2005). The suitability of the marine biotic index (AMBI) to new impact sources along European coasts. Ecological Indicators. 5(1), 19-31. https://doi.org/10.1016/j.ecolind.2004.08.004.

Obst M., Exter K., Allcock A.L., Arvanitidis C.D., Bustamante M., Cancio I., Carreira-Flores D., Chatzinikolaou E., Chatzigeorgiou G., Chrismas N., Clark M.S., Comtet T., Dailianis T., Davies N., Deneudt K., de Cerio O.D., Fortič A., Gerovasileiou V., Hablützel P.I., Keklikoglou K., Kotoulas G., Lasota R., Leite B.R., Loisel S., Lévêque L., Levy L., Malachowicz M., Mavrič B., Meyer C., Mortelmans J., Norkko J., Pade N., Power A.M., Ramšak A., Reiss H., Solbakken J., Staehr P.A., Sundberg P., Thyrring J., Troncoso J.S., Viard F., Wenne R., Zbawicka M., Pavloudi C. (2020). A marine biodiversity observation network for genetic monitoring of hard-bottom communities (ARMS-MBON). Frontiers in Marine Science. 7. https://doi.org/10.3389/fmars.2020.572680.

Pearman J.K., Anlau, H., Irigoie, 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. https://doi.org/10.1016/j.marenvres.2016.04.011.

Pearman J.K., Chust G., Aylagas E., Villarino E., Watson J.R., Chenuil A., Borja A., Cahill A.E., Carugati L., Danovaro R., David R., Irigoien X., Mendibil I., Moncheva S., Rodríguez-Ezpeleta N., Uyarra M.C., Carvalho S. (2020). Pan-regional marine benthic cryptobiome biodiversity patterns revealed by metabarcoding Autonomous Reef Monitoring Structures. Molecular Ecology. 29, 4882–4897. https://doi.org/10.1111/mec.15692.

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. https://doi.org/10.1371/journal.pone.0025026.

Polinski J.M., Bucci J.P., Gasser M., Bodnar A.G. (2019). Metabarcoding assessment of prokaryotic and eukaryotic taxa in sediments from Stellwagen Bank National Marine Sanctuary. Scientific Reports. 9, 14820. https://doi.org/10.1038/s41598-019-51341-3.

Poore G., Wilson G. (1993). Marine species richness. Nature. 361, 597–598. https://doi.org/10.1038/361597a0.

Porter T.M., Hajibabaei M. (2018). Scaling up: A guide to high throughput genomic approaches for biodiversity analysis. Molecular Ecology. 27, 313–38. https://doi.org/10.1111/mec.14478.

Purcell S.W., Mercier A., Conand C., Hamel J.F., Toral-Granda M.V., Lovatelli A., Uthicke, S. (2013). Sea cucumber fisheries: global analysis of stocks, management measures and drivers of overfishing. Fish and Fisheries. 14(1), 34-59. https://doi.org/10.1111/j.1467-2979.2011.00443.x.

Rabalais N.N., Turner R.E., Díaz R.J., Justić D. (2009). Global change and eutrophication of coastal waters. ICES Journal of Marine Science. 66(7), 1528-1537. https://doi.org/10.1093/icesjms/fsp047.

Ransome E., Geller J.B., Timmers M., Leray M., Mahardini A., Sembiring A., Collins A.G., 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, e0175066. https://doi.org/10.1371/journal.pone.0175066.

Roberts C.M., McClean C.J., Veron J.E.N., Hawkins J.P., Allen G.R., McAllister D.E., Mittermeier C.G., Schueler F.W., Spalding M., Wells F., Vynne C., Werner T.B. (2002). Marine biodiversity hotspots and conservation priorities for tropical reefs. Science. 295(5558), 1280–1284. https://doi.org/10.1126/science.1067728.

Robinson C.V., Porter T.M., Wright M.T., Hajibabaei M. (2020). Propylene glycol-based antifreeze as an effective preservative for DNA metabarcoding of benthic arthropods. bioRxiv. https://doi.org/10.1101/2020.02.28.970475.

Roskov Y., Abucay L., Orrell T., Nicolson D., Bailly N., Kirk P.M., Bourgoin T., DeWalt R.E., Decock W., De Wever A. (2018). Species 2000 & ITIS Catalogue of Life, Annual Checklist. Catalogue of Life Checklist. Technical Report.

Rubal M., Costa-Garcia R., Besteiro C., Sousa-Pinto I., Veiga P. (2018). Mollusc diversity associated with the non-indigenous macroalga *Asparagopsis armata* Harvey, 1855 along the Atlantic coast of the Iberian Peninsula. Marine Environmental Research. 136, 1-7. https://doi.org/10.1016/j.marenvres.2018.02.025.

Rubal M., Veiga P., Reis P.A., Bertocci I., Sousa-Pinto I. (2014). Effects of subtle pollution at different levels of biological organization on species-rich assemblages. Environmental Pollution. 191, 101e110. http://dx.doi.org/10.1016/ j.envpol.2014.04.019.

Sawaya N.A., Djurhuus A., Closek C.J., Hepner M., Olesin E., Visser L., Kelbe C., Hubbard K., Breitbart, M. (2019). Assessing eukaryotic biodiversity in the Florida Keys National Marine Sanctuary through environmental DNA metabarcoding. Ecology and Evolution. 9(3), 1029-1040. https://doi.org/10.1002/ece3.4742.

Sedano F., Florido M., Rallis I., Espinosa F., Gerovasileiou V. (2019). Comparing sessile benthos on shallow artificial versus natural hard substrates in the Eastern Mediterranean Sea. Mediterranean Marine Science. 20(4), 688–702. https://doi.org/10.12681/mms.17897.

Sedano F., Navarro-Barranco C., Guerra-García J.M., Espinosa F. (2020). Understanding the effects of coastal defence structures on marine biota: The role of substrate composition and roughness in structuring sessile, macro-and meiofaunal communities. Marine Pollution Bulletin. 157, 111334. https://doi.org/10.1016/j.marpolbul.2020.111334.

Shokralla S., Porter T.M., Gibson J.F., Dobosz R., Janzen D.H., Hallwachs W., Golding G.B., Hajibabaei M. (2015). Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform. Scientific Reports. 5, 9687. https://doi.org/10.1038/srep09687.

Shokralla S., Spall J.L., Gibson J.F., Hajibabaei M. (2012). Next-generation sequencing technologies for environmental DNA research. Molecular Ecology. 21, 1794–1805. https://doi.org/10.1111/j.1365-294X.2012.05538.x.

Shum P., Barney B.T., O'Leary J.K., Palumbi S.R. (2019). Cobble community DNA as a tool to monitor patterns of biodiversity within kelp forest ecosystems. Molecular Ecology Resources. 19, 1470–1485. https://doi.org/10.1111/1755-0998.13067.

Simberloff D., Martin J.L., Genovesi P., Maris V., Wardle D., Aronson J., Courchamp F., Galil B., Garcia-Berthou E., Pascal M., Pyšek P., Sousa R., Tabacchi E., Vilà M. (2013). Impacts of biological invasions: what's what and the way forward. Trends in Ecology and Evolution. 28, 58–66. https://doi.org/10.1016/j.tree.2012.07.013.

Solan M., Cardinale B.J., Downing A.L., Engelhardt K.A., Ruesink J.L., Srivastava D.S. (2004). Extinction and ecosystem function in the marine benthos. Science. 306(5699), 1177-1180. https://doi.org/10.1126/science.1103960.

Taberlet P., Coissac E., Pompanon F., Brochmann C., Willerslev E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. Molecular Ecology. 21, 2045–2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x.

Tang C.Q., Leasi F., Obertegger U., Kieneke A., Barraclough T.G., 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. 109(40), 16208-16212. https://doi.org/10.1073/pnas.1209160109.

Teixeira H., Salas F., Borja Á., Neto J.M., Marques J.C. (2008). A benthic perspective in assessing the ecological status of estuaries: The case of the Mondego estuary (Portugal). Ecological Indicators. 8, 404–416. https://doi.org/10.1016/j.ecolind.2007.02.008.

Templado J., Paulay G., Gittenberger A., Meyer C. (2010). Sampling the marine realm. ABC Taxa. 8, 273-307.

Tittensor D.P., Mora C., Jetz W., Lotze H.K., Ricard D., Vanden B.E., Worm B. (2010). Global patterns and predictors of marine biodiversity across taxa. Nature. 466(7310), 1098–1011. https://doi.org/10.1038/nature09329.

Turon X., Palac C., Pr K., Wangensteen O.S. (2020). From metabarcoding to metaphylogeography: separating the wheat from the chaff. Ecological Applications. 30, 19. https://doi.org/10.1002/eap.2036.

van der Loos L.M., Nijland R. (2020). Biases in bulk: DNA metabarcoding of marine communities and the methodology involved. Molecular Ecology. https://doi.org/10.1111/mec.15592.

van Hoey G., Borja A., Birchenough S., Buhl-Mortensen L., Degraer S., Fleischer D., Kerckhof F., Magni P., Muxika I., Reiss H., Schröder A., Zettler M.L. (2010). The use of benthic indicators in Europe: from the Water Framework Directive to the Marine Strategy Framework Directive. Marine Pollution Bulletin. 60(12), 2187-2196. https://doi.org/10.1016/j.marpolbul.2010.09.015.

Veiga P., Rubal M., Cacabelos E., Maldonado C., Sousa-Pinto I. (2014). Spatial variability of macrobenthic zonation on exposed sandy beaches. Journal of Sea Research. 90, 1-9. http://dx.doi.org/10.1016/j.seares.2014.02.009.

Veiga P., Torres A.C., Aneiros F., Sousa-Pinto I., Troncoso J.S., Rubal M. (2016). Consistent patterns of variation in macrobenthic assemblages and environmental variables over multiple spatial scales using taxonomic and functional approaches. Marine Environmental Research. 120, 191-201. http://dx.doi.org/10.1016/j.marenvres.2016.08.011.

Vieira P.E., Desiderato A., Holdich D.M., Soares P., Creer S., Carvalho G.R., Costa F.O., Queiroga H. (2019). Deep segregation in the open ocean: Macaronesia as an evolutionary hotspot for low dispersal marine invertebrates. Molecular Ecology. 28(7), 1784-1800. https://doi.org/10.1111/mec.15052.

Wangensteen O.S., Cebrian E., Palacín C., Turon X. (2018a). Under the canopy: Community-wide effects of invasive algae in Marine Protected Areas revealed by metabarcoding. Marine Pollution Bulletin. 127, 54–66. https://doi.org/10.1016/j.marpolbul.2017.11.033.

Wangensteen O.S., Palacín C., Guardiola M., Turon X. (2018b). DNA metabarcoding of littoral hard-bottom communities: high diversity and database gaps revealed by two molecular markers. PeerJ. 6, e4705. https://doi.org/10.7717/peerj.4705.

Weigand A.M., Desquiotz N., Weigand H., Szucsich N. (2021). Application of propylene glycol in DNA-based studies of invertebrates. Metabarcoding and Metagenomics. 5, e57278. https://doi.org/10.3897/mbmg.5.57278.

Weigand H., Beermann A.J., Čiampor F., Costa F.O., Csabai Z., Duarte S., Geiger M.F., Grabowski M., Rimet F., Rulik B., Strand M., Szucsich N., Weigand A.M., Willassen E., Wyler S.A., Bouchez A., Borja A., Čiamporová-Zaťovičová Z., Ferreira S., Dijkstra K.B., Eisendle U., Freyhof J., Gadawski P., Graf W., Haegerbaeumer A., van der Hoorn B.B., Japoshvili B., Keresztes L., Keskin E., Lesse F., Macher J.N., Mamos T., Paz G., Pešić V., Pfannkuchen D.M., Pfannkuchen M.A., Price B.W., Rinkevich B., Teixeira M.A.L., Várbíró G., Ekrem T. (2019). DNA barcode reference libraries for the monitoring of aquatic biota in Europe: Gap-analysis and recommendations for future work. Science of the Total Environment. 678, 499–524. http://doi.org/10.1016/j.scitotenv.2019.04.247.

Xie Y., Zhang X., Yang J., Kim S., Hong S., Giesy J.P., Yim U.H., Shim W.J., Yu H., Khim J.S. (2018). eDNA-based bioassessment of coastal sediments impacted by an oil spill. Environmental Pollution. 238, 739–748. http://doi.org/10.1016/j.envpol.2018.02.081.

Yang C., Tan S., Meng G., Bourne D.G., O'Brien P.A., Xu J., Liao S., Chen A., Chen X., Liu S. (2018). Access COI barcode efficiently using high throughput Single-End 400 bp sequencing. BioRxiv. 498618. https://doi.org/10.1101/498618.

Yeo D., Srivathsan A., Meier R. (2020). Longer is not always better: Optimizing barcode length for large-scale species discovery and identification. Systematic Biology. https://doi.org/10.1093/sysbio/syaa014.

Yu D.W., Ji Y., Emerson B.C., Wang X., Ye C., Yang C., Ding Z. (2012). Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. Methods in Ecology and Evolution. 3, 613–623. https://doi.org/10.1111/j.2041-210X.2012.00198.x.

Zinger L., Bonin A., Alsos I.G., Bálint M., Bik H., Boyer F., Chariton A.A., Creer S., Coissac E., Deagle B.E., De Barba M., Dickie I.A., Dumbrell A.J., Ficetola G.F., Fierer N., Fumagalli L., Gilbert M.T.P., Jarman S., Jumpponen A.,

Kauserud H., Orlando L., Pansu J., Pawlowski J., Tedersoo L., Thomsen P.F., Willerslev E., Taberlet P. (2019). DNA metabarcoding—Need for robust experimental designs to draw sound ecological conclusions. Molecular Ecology. 28, 1857–1862. https://doi.org/10.1111/mec.15060.

Zizka V.M.A., Leese F., Peinert B., Geiger M.F. (2019). DNA metabarcoding from sample fixative as a quick and voucher-preserving biodiversity assessment method. Genome. 62, 122–136. https://doi.org/10.1139/gen-2018-0048.

Chapter 2

Gap-analysis and annotated reference library for supporting macroinvertebrate metabarcoding in Atlantic Iberia

Abstract

DNA metabarcoding provides a rapid and effective identification tool of macroinvertebrate species. The accuracy of species-level assignment, and consequent taxonomic coverage, relies on comprehensive DNA barcode reference libraries, which, due to incompleteness, are currently a recognized limitation for metabarcoding applications. In this study, we assembled a comprehensive reference library of DNA barcodes for Atlantic Iberia marine macroinvertebrate species, assessed gaps in species coverage and examined data ambiguities. Initially, an Iberian species checklist for the three dominant groups of marine macroinvertebrates was compiled, comprising 2827 species (926 Annelida, 638 Crustacea and 1263 Mollusca). A total of 18162 DNA sequences of the cytochrome c oxidase subunit I barcode region (COI-5P) matching the species checklist were compiled in a BOLD dataset, where taxonomic discordances were evaluated and cases of deep intraspecific divergence flagged. Gap-analysis showed that 63% of the Iberian macroinvertebrate species still lack a DNA barcode. Coverage gaps varied considerably across taxonomic groups with Mollusca displaying the highest sequence representation in the dataset (427 species, 49% of the total number of sequences), and Crustacea the highest species coverage with 338 species barcoded (53% of the checklist). In contrast, Polychaeta displayed the lower levels of completion (288 species, 16% of the total number of sequences). In total, 1545 Barcode Index Numbers (BINs) were assigned to 1053 barcoded species, of which 66% were taxonomically concordant, 26% displayed multiple BINs and 8% were discordant. Overall, results show that there is still a large portion of marine invertebrate taxa in this region of Europe pending barcode coverage, even considering only the dominant groups. However, the most notable finding was the relevant proportion of species flagged for significant intraspecific divergence and possible hidden diversity. The annotated reference library and gap-analysis here provided can therefore contribute to prioritize marine macroinvertebrate taxa for future research efforts and barcode coverage.

Keywords: Macroinvertebrates; DNA barcoding; Reference library; Cytochrome c oxidase subunit I; Atlantic Iberia.

2.1 Introduction

The Atlantic coast of the Iberian Peninsula occupies a central geographic position in the North east Atlantic, spreading along intermediate latitudes (between 43° 47'N and 36° 00'N), connecting the north temperate and warm sub-tropical waters (Spalding et al., 2007; Briggs and Bowen, 2012). This Atlantic region is right at the core of the Lusitanian biogeographic province (Spalding et al., 2007; Briggs and Bowen, 2012) that harbors a diverse marine fauna, enriched by the faunas from the various adjacent regions, such as the Mediterranean, the sub-tropics and Macaronesia, as well as the faunas from further north and western Atlantic. Many invertebrate species have their northern or southern range limits in this area (e.g. Boaventura et al., 2002; Pereira et al., 2006; Lima et al., 2007), which makes it a particularly suitable region to monitor the impact of global change in marine species ranges in the NE Atlantic. Furthermore, offshore Atlantic Iberia waters constitute one of the largest routes of the globe for maritime traffic (Nunes et al., 2014; Pejovic et al., 2016), which, together with major commercial ports and numerous recreational marinas in Portugal and Spain, make this region highly susceptible to exposure to non-indigenous marine species (Chainho et al., 2015; Rubal et al., 2018). Monitoring of coastal fauna in the Iberian Atlantic Peninsula is therefore of prime importance for early assessment of impacts and changes in marine communities and ecosystems that may have repercussions in other areas of the NE Atlantic (Araújo et al., 2009; Pascual et al., 2010; Miralles et al., 2016; Múrria et al., 2019).

Due to their rapid and sensitive response to environmental and human pressures, marine macroinvertebrates have been widely used as bioindicators of ecological status in marine ecosystems (Aylagas et al., 2014). Up to now, long term monitoring of coastal ecosystems in the Iberian Atlantic Peninsula has been carried out through morphology-based community assessments (e.g. Guerra-García et al., 2011; Veiga et al., 2016), including monitoring of macroinvertebrate assemblages in the scope of EU's Water Framework Directive (WFD, Directive 2000/60/EC) and Marine Strategy Framework Directive (MSFD, Directive 2008/56/EC). Morphology-based diagnosis is particularly challenging in Atlantic Iberia, considering the diversity of faunal assemblages, including species coming from various adjacent regions, and the lack of dedicated taxonomic keys for this region (especially when compared with other marine regions such as Mediterranean – Ruffo, 1982; the British Isles – Lincoln, 1979; Naylor 1992; or Northwest Europe – Hayward and Ryland, 1995). Moreover, reports displaying growing evidence for the existence of a sizeable proportion of hidden or cryptic diversity among marine invertebrates (Hupalo et al., 2018;

Teixeira et al., 2019; Vieira et al., 2019), including many taxa that occur in this region (e.g. Borges et al., 2016; Lobo et al., 2016, 2017), further call into question the accuracy of morphology-based assessments.

The recent development of DNA metabarcoding approaches for species assessments (Hajibabaei et al., 2012; Cristescu, 2014) provides an opportunity to complement morphology-based procedures, thereby globally improving the accuracy, throughput and efficiency of marine monitoring, including macroinvertebrate communities (Bourlat et al., 2013; Cowart et al., 2015; Aylagas et al., 2018; Pearman et al., 2018). In addition to constituting the single available method to diagnose cryptic species (Lindeque et al., 2013), strengths of DNA-based approaches include reduced ambiguity and greater accuracy, identification of small taxa, immature or larval stages, and possibility of direct comparison among sites and studies and future verification of the identifications (Leese et al., 2016, 2018). It also enables higher spatial-temporal frequency in monitoring due to higher throughput (Bush et al., 2019). However, the usefulness and efficiency of DNA metabarcoding depends heavily on the extent of the taxonomic coverage and the quality of records available in the reference libraries of DNA barcodes that underpin the method (Sidall et al., 2009; Leray and Knowlton, 2016). The development of several biomonitoring programs associated with high-throughput biodiversity data prompts the necessity to provide quality assurance for DNA barcodes (Leese et al., 2016, 2018; Oliveira et al., 2016; Weigand et al., 2019). In fact, important taxonomic gaps in the reference libraries of DNA barcodes of marine invertebrates have been recently reported for the European marine regions, which are typically much larger than their freshwater counterparts. Some records are flagged as doubtful barcodes and inadequate quality standardization of reference barcodes can affect the reliability of a reference library. This includes identification errors, sequence contamination, incomplete reference data without trace files or primer information and inadequate data management (Weigand et al., 2019).

Considering the above described relevance of the Atlantic Iberia region and the importance of up-to-date DNA-based technologies to support macroinvertebrate monitoring in this region, we conducted a comprehensive assessment of the gaps in the regional reference library of COI barcodes for the three most prominent coastal marine taxa (Annelida, Crustacea and Mollusca) occurring in this area. We also reviewed the taxonomic congruency of the COI barcode records and provided a comprehensive and annotated reference library for the target taxa occurring in Atlantic Iberia.

2.2 Material and methods

2.2.1 Study area and checklist

A comprehensive species-level checklist of marine macroinvertebrate species occurring in Iberian Atlantic Coast was compiled using the World Register of Marine Species (WoRMS) database (http://www.marinespecies.org/) and literature records (for checklist details with species information consult Table S.1[1]; to consult references see Table S.2). The study area comprised the marine region of Continental Atlantic Iberia, i.e., between the France-Spain Atlantic border to the strait of Gibraltar (Macaronesia not included). We assessed only selected taxonomic ranks among the three most dominant groups of marine macroinvertebrates: Crustacea (Malacostraca: Amphipoda, Decapoda and Isopoda; Thecostraca: Balanomorpha), Annelida (Polychaeta) and Mollusca (Bivalvia, Gastropoda and Polyplacophora). The validity of the species names in the final checklist and their assignment as "marine" was verified in WoRMS database with the package “worms” (Holstein, 2018), through the software R 3.6.1 (R core Team, 2019; www.r-project.org).

2.2.2 Data mining and BOLD dataset creation

All the available COI-5P sequences matching the species names in the checklist for Atlantic Iberia were mined from the Barcode of Life Data system (BOLD; Ratnasingham and Hebert, 2007) using the R package “bold” (Chamberlain, 2019). Records without information on species name, containing COI sequences with less than 500 base pairs and flagged for contamination, stop-codons or indels were subsequently removed. To this dataset, we added new sequences of specimens collected at the Iberian Atlantic coast (dataset DS-AIMARINV, which also includes records obtained by us from our past publications). A final dedicated dataset which aggregate all compiled DNA barcodes (DS-GAIMARIN - doi.org/10.5883/DS-GAIMARIN; Table 2.1) was created in BOLD.

The new COI barcode sequences were obtained from specimens collected on the Portuguese coast, following published protocols (Borges et al., 2016; Lobo et al., 2016, 2017), and were submitted to GenBank (accessions and specimen list are available in Table S.3).

[1] The supplementary material of this chapter is available online at https://doi.org/10.1016/j.rsma.2020.101307.

Table 2.1. Number of COI-5P sequences generated under this study and number of COI-5P sequences retrieved from BOLD, compiled in the dataset DS-GAIMARIN for each target taxa, with the associated BOLD project code.

Target taxa	BOLD Project code	Number of COI-5P sequences
Crustacea: Decapoda	MLALE*	22
Crustacea: Isopoda	ISOAL*; WBEC*	3
Crustacea: Balanomorpha	FCCOM*	27
Annelida: Polychaeta	PCALE*	9
Mollusca: Bivalvia	BIPM*; BIV*; BVALN*; METP*	39
Mollusca: Gastropoda	GTALE*	7
Mollusca: Polyplacophora	PIPM*	8
Crustacea, Annelida and Mollusca [a]		18047
	Total	18162

*New data generated on this study. [a] Retrieved from Bold

2.2.3 Data processing and analyses

To conduct a global gap-analysis of the barcoded species from Atlantic Iberia, we compared the species checklist with all publicly available COI-5P sequence records in BOLD. A species was considered successfully barcoded if at least one COI-5P sequence (>500bp) was available. The geographic origin of the specimens was also recorded. All the records sampled between the co-ordinates lat 42.00 and 44.00 and between long -11.00 and -02.00 (North Continental Atlantic Iberian Peninsula) or between lat 36.00 and 42.00 and between long -11.00 and -05.30 (South Continental Atlantic Iberian Peninsula) or with clear indication that were sampled in Continental Atlantic Iberia were considered as "Atlantic Iberian Peninsula". All the records with clear country information outside from the delimited area were considered as "not Atlantic Iberian Peninsula". Ambiguous records indicating Iberian Peninsula or Atlantic Ocean (i.e., doubts if they were sampled in Atlantic Continental Iberia or elsewhere) were considered "unknown". For the purpose of data analyses, we first considered records of species from Atlantic Iberian Peninsula, second

selected species with records indicating geographic sampling collection outside Iberian Peninsula; and lastly the "unknown" records.

All the species in the dataset which were assigned to a Barcode Index Number (BIN) (Ratnasingham and Hebert, 2013) were annotated with one of three possible taxonomic congruency grades: discordant (i.e. more than one species assigned to the same BIN), multiple BINs (i.e., one species assigned to more than one BIN) and concordant (i.e., one species assigned to only one BIN). The BINs assigned to different species (i.e. discordant BINs) were carefully inspected by checking their placement in NJ phenograms, looking for the valid species names, synonyms or contaminations, and by inspecting BINs' content on BOLD database. Namely, to further verify the congruency between BINs and morphospecies, neighbor joining (NJ) phenograms for each phylum, were built in MEGA v7, using Kimura-2-parameter (K2P) substitution model (Kimura, 1980). Node support was assessed through 1000 bootstrap replicates. Only three selected sequences from each BIN were used to construct the NJ phenograms. The selection was performed by using the following criteria in the same order: 1) without "N" and without gaps and whenever possible, select sequences with 658 bp; 2) sequences between 650 bp and 657 bp; 3) sequences between 600 bp and 649 bp; 4) sequences between 500 bp and 599 bp; 5) sequences higher than 658 bp; 6) any available sequences. When more than three sequences were compliant with the above criteria, only three were randomly selected. In the case of species without an attributed BIN, we selected three sequences from each species following the above criteria. All sequences were aligned using MAFFT v7 (https://mafft.cbrc.jp/alignment/server/; Katoh & Standley, 2013).

The bioinformatic pipeline developed to carry out these analyses is available at https://github.com/pedroemanuelvieira/Iberian-Peninsula-DNA-Reference-Library.

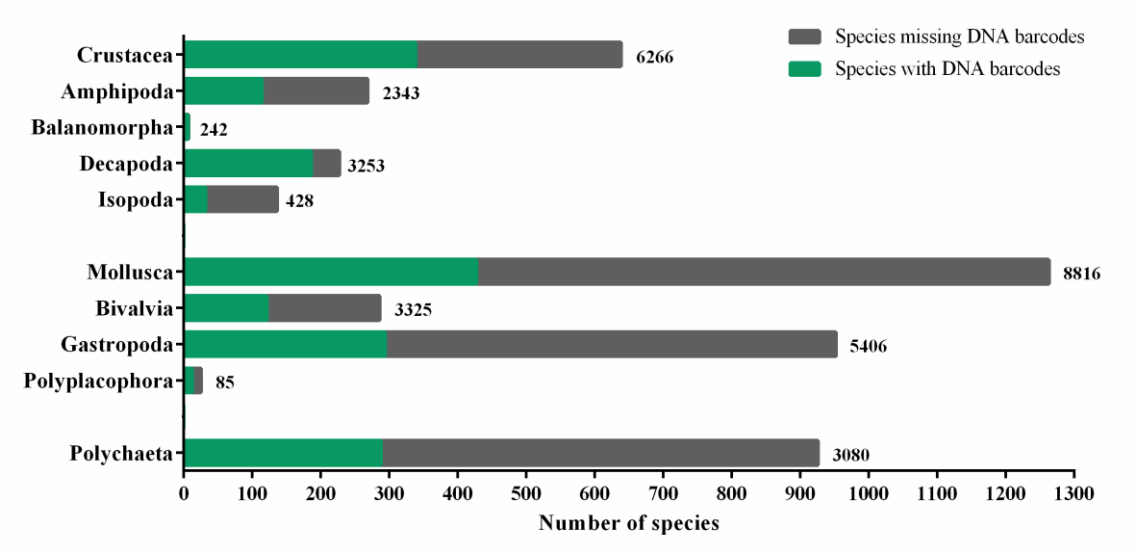
2.3. Results

The final checklist had a total of 2827 marine macroinvertebrate species occurring in the Atlantic Iberia, belonging to three major groups (926 Annelida, 638 Crustacea, and 1263 Mollusca). The distribution by taxonomic groups, the number of sequences per group and the geographic region of specimen collection are displayed in Figure 1 (see detailed information in Table S.3 and Table S.4).

The DS-GAIMARIN dataset is composed by 18162 COI-5P sequences belonging to 1053 species, assigned to 1545 BINs. One-hundred and fifteen new DNA barcodes were generated under this study, among which two species were barcoded for the first time, namely the decapod *Gilvossius tyrrhenus* (Petagna, 1792) and the polyplacophora *Leptochiton albemarlensis* A. G. Smith & Ferreira, 1977.

Mollusca was the most well represented taxon in the dataset in number of sequences (8816 sequences, 49%), and the most diverse class, in terms of species, was Gastropoda (952 species), which also displayed the highest number of sequences (5406). Crustacea also had a high proportion of species with DNA barcodes (6266 sequences, 35% of sequence representation), for which Amphipoda was the most well represented order in terms of species (268 species), although Decapoda recorded a higher number of DNA barcode sequences (3253 sequences, 52% of sequence representation). On the other hand, Polychaeta displayed very low numbers of sequences (16%), although well represented with species in the list (approximately 33% of the total number of species).

A.



B.

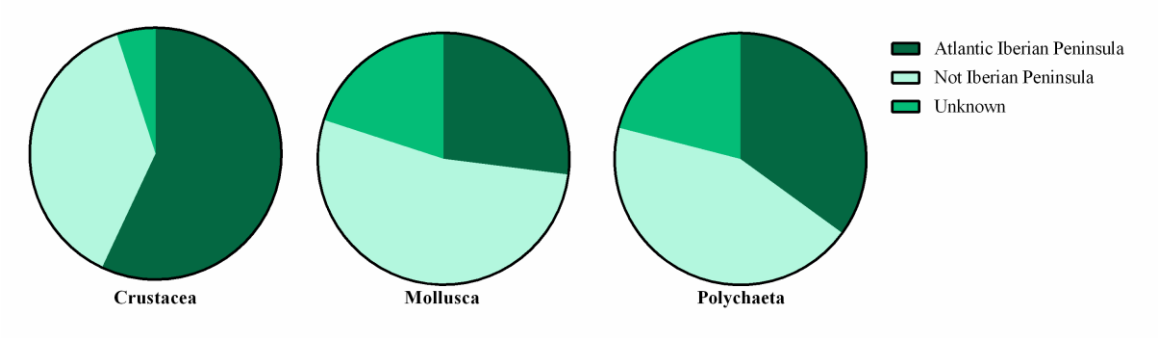


Figure 2.1. A. DNA barcode coverage for marine macroinvertebrate species occurring in Atlantic Iberia. Barcode coverage with at least one COI-5P sequence per species (green bar). Numbers on the right side of each bar refer to the total number of sequences. **B.** Partioning of the geographic origin of specimens available in the reference library: samples in Atlantic Iberia, outside of Iberian Peninsula and uncertain geographical information (unknown).

The availability of DNA barcodes for the examined taxa varied considerably across taxonomic groups (Fig. 2.1. A), and in total only 37% (1053 species) of the species had at least one barcode sequence deposited in BOLD. Among the three selected groups, Polychaeta had the lowest barcode coverage, with only 31% (corresponding to 288 species) of the total species represented in the checklist being barcoded. In Mollusca 34% of the species were barcoded, however among the three major classes more than 50% were still missing DNA barcode sequences. Despite Gastropoda having the highest number of representative sequences it displayed a lower level of completion (31%), than Bivalvia (42%) or Polyplacophora (52%). Yet

the number of listed taxa is highly disparate for these classes: 952 for Gastropoda, 286 for Bivalvia and 25 for Polyplacophora. Overall, Crustacea had the largest coverage with 53% (338 species barcoded), but Decapoda and Balanomorpha reach more than 80% of total species barcoded, while Amphipoda and Isopoda displayed very low completion; 43% and 24%, respectively.

Although Crustacea did not display the highest number of sequences, a detailed analysis of the geographic region from where the specimens were sampled reveals that the highest representation of species from Atlantic Iberia was found for this group (57% of total number of species, Fig. 2.1. B). The results also showed that the majority of the sequences had the specimens sampling collection information associated, and for only 16% of barcoded species there was no data or insufficient data (i.e. "unknown"). Sixty-seven species in the list are non-indigenous for the Iberian Peninsula, of which 61% (41 species) had at least one barcode sequence deposited in BOLD, while 39% (26 species) lack a barcode sequence (Table S.5.).

Overall, the majority of BINs were considered concordant (i.e. one BIN = one species): 649 species corresponding to 42% of the total number of BINs (Fig. 2.2). A total of 284 species were assigned to more than one BIN (corresponding to 831 BINs, 27% of the species), and Mollusca displayed the highest percentage of species assigned to multiple BINs (124 species, 43% of BINs). Among Mollusca, Gastropoda had the highest number of species displaying multiple BINs (80 species, 59% of BINs), however it was also the class harboring more species and consequently more representative sequences. Only 11% of the total number of BINs were discordant (i.e. 120 BINs were shared by more than one species). Gastropods showed the highest levels of discordance (36 species), whereas Polyplacophora, Isopoda and Balanomorpha did not display any discordant BINs. However, a subsequent inspection of BINs revealed an overestimation and unrealistic percentage of discordant BINs. Following a careful inspection, 33 discordant BINs displayed concordance or can be assigned to other species, mainly due to misidentifications. Therefore, the number of discordant BINs decreased to 87 species (8%; Table S.6), and consequently the number of concordant and multiple BINs increased to 676 species (66%) and 259 species (26%; Table S.7), respectively.

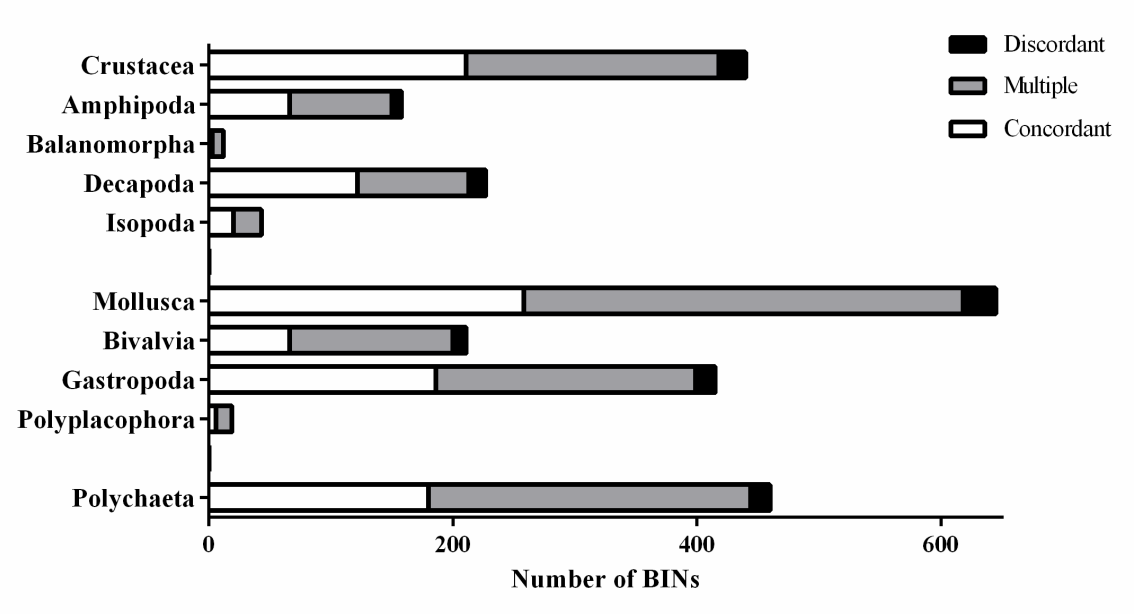


Figure 2.2. Number of BINs according to taxonomic congruence annotations (concordant, multiple or discordant) for each taxonomic group on the Atlantic Iberia reference dataset.

Phylogenetic trees were constructed for taxonomic reliability inspection (Figures S.1-S.8). A total of 3178 COI-5P sequences distributed over the three taxonomic groups (1010 Crustacea, 1343 Mollusca, and 825 Polychaeta) were used to construct the trees. The number of species represented by only one sequence per BIN (singletons) was 65 for Crustacea, 103 for Mollusca and 127 for Polychaeta. Furthermore, 170 species (61 Crustacea, 80 Mollusca and 29 Polychaeta) displayed a high intraspecific divergence, and the groups with the highest values were Gastropoda (46 species) and Decapoda (32 species), followed by Bivalvia (29 species) and Amphipoda (20 species).

2.4 Discussion

The current study highlights three main considerations: first, reference libraries still lack representative barcodes for many marine macroinvertebrate species belonging to dominant faunal groups; second, a considerable number of species apparently integrate hidden or undescribed diversity; and third, a comparatively low proportion of taxonomic incongruences were detected, which may eventually impact the accuracy of current DNA-based assessment and biomonitoring of marine ecosystems, though we partially sorted them out by auditing and annotating our compiled reference library.

Marine macroinvertebrates are among the most phylogenetically diverse communities, thereby constituting a particularly demanding component for morphology-based biomonitoring, and also a harder target to achieve a comprehensive DNA barcode reference library (Lobo et al., 2017). Yet, DNA metabarcoding's taxonomic span of detection and degree of accuracy is highly depend on reference libraries completion and reliability of the records. In the current study the gap of DNA barcodes found was considerably high for the examined taxonomic groups (63%). This result was not unexpected since a number of studies already revealed a high prevalence of gaps of DNA barcodes for specific taxa (e.g. Aylagas et al., 2014; Abad et al., 2016; Lobo et al., 2016), and other studies showed disagreements between molecular and morphological assignments which are mostly associated to incompleteness of reference sequences databases (Kelly et al., 2017; Weigand et al., 2019). Furthermore, marine biodiversity assessment is challenging due to geographical large-scale sampling effort, which has a critical impact on species assessment (Bergsten et al., 2012). These factors have negative impacts on taxonomic research, which leads to a higher proportion of undescribed diversity and affect the outcome of richness of a community (Pawlowski et al., 2018).

Our results are comparable to those obtained with other checklists compiled for marine species: the AZTI Marine Biotic Index (AMBI) list (Borja et al., 2000; Aylagas et al., 2014; Weigand et al., 2019) and the European Register of Marine Species (ERMS) list (Weigand et al., 2019), both targeting European marine occurring taxa, but that differ in the taxonomic composition. Although for the three taxonomic groups our checklist had similar number of taxa as the AMBI list (2827 and 2560, respectively) and much lesser than the ERMS list (14207 species), the gap found in the current study was more similar to the one found previously for the ERMS' list (70% for the ERMS vs 63% for ours vs 50% for AMBI) (Weigand et al., 2019).

The number of DNA barcodes available on public databases can be somehow related with the number of dedicated studies, and consequently barcoding projects associated (Weigand et al., 2019). For example, many projects and studies were dedicated to complete the reference sequences databases for fishes (Costa et al., 2012; Keskin and Atan, 2013; Oliveira et al., 2016; Cariani et al., 2017), while the number of macroinvertebrate barcoding dedicated projects are much lower. Indeed, the obtained differences on DNA barcode completion among the three taxonomic groups can be explained by the frequency that specific taxonomic groups and/or species are targeted in barcoding studies (Barco et al., 2016; Lobo et al., 2016). We found many barcoding-based studies dedicated to crustaceans (Costa et al., 2007; Matzen da Silva et al., 2013; Raupach et al., 2015; Lobo et al., 2017), which can highly increase the representativeness of

sequences belonging to this group in genetic databases, and is probably the best explanation for the largest coverage of Crustacea found in the current study. Furthermore, most of these studies based on crustaceans were developed in Atlantic Iberia, which can explain the highest representation of this group with records from Atlantic Iberia. Our results reveal the need to increase the projects and studies dedicated to marine macroinvertebrate species, in particular for Annelida and Mollusca. However, while reference libraries are far from being complete, the generation of DNA barcodes for the most frequent species occurring at a particular site or region may overcome databases incompleteness and consequently improve taxonomic assignment using DNA-based tools in local studies (Ayalagas et al., 2014; Abad et al., 2016). Although we registered a relatively low proportion of sequences (16%) with no geographic data or insufficient data (i.e. "unknown" sequences represented on Fig. 2.1. B.), it is important to stress the relevance of metadata in public databases, particularly the geographic origin of the specimens, which is especially critical for a library still with considerable gaps and numerous poorly-represented species.

Globally, the comparative analysis between morphology-based taxonomic identification and BINs assignments exposed a sizeable amount of discordances. Notably, it also suggested that species diversity assessed through morphology can be currently underestimated by as much as 50% of the target taxa, with suspected hidden diversity recorded on over 5% of the examined morphospecies. These findings are transversal to all taxonomic groups and can be explained by different reasons. In some cases, the species names have not yet been updated in the BOLD database, and sequences misidentified combined with taxa absent in the databases will generate incorrect taxonomic identifications. For example, two specimens of gastropods were morphologically identified to genus-level as *Nassarius* sp. and *Ocenebrina* sp. However, based on barcodes and phylogenetic trees construction, both cases can be now identified as *Tritia incrassata* (Strøm, 1768) and *Ocenebra edwardsii* (Payraudeau, 1826). Moreover, two decapods *Melicertus kerathurus* (Forskål, 1775) and *Penaeus kerathurus* (Forskål, 1775), were attributed to the same BIN (BOLD:AAB4142), but a confirmation on WoRMS of the taxonomic status of the scientific names revealed that *M. kerathurus* is currently unaccepted and was updated to *P. kerathurus*. A closer look to the phylogenetic tree generated in the current study, suggested that other BINs discordances can be related to misidentifications. For example, the BIN BOLD:AAW8076 had sequences identified as *Caprella acanthifera* Leach, 1814 and *Caprella danilevskii* Czerniavski, 1868, which grouped on the same clade with low divergence. However *C. danilevskii* shall be the correct taxonomic identification, since there is another BIN (BOLD:AAY5434) identified as *C. acanthifera* grouped in another clade which is recognized as

a species-complex since 1998 (Krapp-Schickel and Vader, 1998). Overall, our careful inspection of the composition of the discordant BINs revealed that most were related to misidentifications or synonyms.

Consistently, all taxonomic groups analyzed displayed a fair amount of cases of high intraspecific divergence, probably related with hidden or cryptic diversity, of which most of them were already reported in previous studies (Best and Stachowicz, 2013; Layton et al., 2014; Leray and Knowlton, 2015; Trickey et al., 2016; McCarthy et al., 2019; Teixeira et al., 2019; Vieira et al., 2019). For example, for the gastropod genus *Doto* represented in the reference library by 8 distinct species, has been pointed out as an extremely challenging group for taxonomic identifications due to their small body size, similar color patterns and lack of distinctive morphological characters (Morrow et al., 1992, Pola and Gosliner, 2010). In our results, *Doto coronata* (Gmelin, 1791) and *Doto koenneckeri* Lemche, 1976 were each one assigned to two different BINs and grouped in different clades with high divergence (>8%). Some researchers described these lineages as a complex (Korshunova et al., 2016; Shipmand and Gosliner, 2015), but more taxonomic and molecular work are still needed to solve this issue. Another example of an observed cryptic complex is the polychaete *Syllis gracilis* Grube, 1840. DNA barcodes for this species were sorted into multiple lineages with an unbalanced representation: 34 sequences assigned to 11 BINs. This cryptic complex has been already disclosed, however a combination of different interactions among environmental features and biogeographical factors have been hindering its full interpretation (Langeneck et al., 2019).

Although previous studies on Amphipoda revealed a majority of monophyletic clades consistent with consolidated morphospecies (Raupach et al., 2015), there is still considerable taxonomic instability in particular species, which display among the highest levels of intraspecific divergence here recorded. This is the case, for example of the *Apohyale stebbingi* Chevreux, 1888 complex (Desiderato et al., 2019), which was assigned to two different BINs (BOLD:AAI8298 and BOLD:ACX2700) diverging over 13% K2P.

One of the salient advantages of metabarcoding compared to morphology-based monitoring is the ability to detect and document the occurrence of cryptic species. However, because metabarcoding procedures typically use shorter fragments then the full COI-5P, one question that may arise is if there will be still enough resolution to discriminate cryptic species in such conditions. At least for COI metabarcoding of marine invertebrates, studies have shown that there is very little loss of discrimination ability for segments of COI-5P down to 200 base pairs (Hollatz et al., 2017). However, this may not hold for other markers, such as for example 18S rDNA sequences, which have been documented to have little discrimination ability at the species-level (Tang et al., 2012; Lejzerowicz et al., 2015; Danovaro et al., 2016). On the other hand,

instances of hybridization or mitochondrial introgression between closely related species will fail detection by metabarcoding (like through regular DNA barcoding), but such phenomena will be very likely overlooked by morphology-based monitoring as well (Cowart et al., 2015; Pawlowski et al., 2018). There are of intrinsic pitfalls both in morphology- and barcode-based identifications. Thus, the combination of morphological identification with DNA barcoding in an integrative approach for monitoring biodiversity contribute to significantly facilitate comparative studies of genetic diversity in different species. In addition, this integrative approach also facilitates comprehensive analyses of a given taxonomic assemblage and provides insights into the patterns of genomic diversity within species.

The addition of publicly available sequences obtained from specimens collected in other geographic regions allowed to understand patterns of concordance/discordance between BINs. For example, the polychaete *Cirriformia tentaculata* (Montagu, 1808) was assigned to two distinct BINs which grouped in two clades: BOLD:ACI3598 corresponding to samples originated from China, and BOLD:ACI2312 corresponding to samples collected from Portugal. This can be an evidence of possible cryptic polymorphism in this species, already pointed out in a previous study (George, 1967). However, BINs were composed by a low number of sequences (less than 3 sequences), which are not enough to reach a strong conclusion.

Considering the diversity of faunal assemblages combined with the introduction of non-indigenous Species (NIS) in the Iberian Atlantic coast, it is especially important to early assess and monitor the impacts and changes in marine species range, identifying possible biological invasions and enable the development of mitigation strategies (Briski et al., 2016; Rey et al., 2019; Viard et al., 2019). In order to use metabarcoding as a tool to early detect and improve monitoring of NIS in coastal and marine ecosystems, it is extremely important to complete the number of missing barcode sequences for NIS (Briski et al., 2016; Ardura, 2019), as well as to solve problems of multiple or discordant BINs associated to NIS, since in this case species-level identification is mandatory and wrong identifications can trigger action or inaction when not required.

The detection of a reasonable number of marine macroinvertebrates still missing DNA barcodes and the presence of hidden or undescribed diversity in the reference library compiled in this study, highlight the urgent need to complete and curate reference sequences databases for such important marine groups. The reference library compiled, audited and annotated in the current study is ought to be a valuable support to improve the precision of taxonomic assignments in metabarcoding studies in Atlantic Iberia and to overcome under- or overestimation of species richness.

2.5 Conclusions

The reference library compiled in this study covers the most dominant groups for marine macroinvertebrate species occurring in the Iberian Atlantic coast, which are the most commonly used species in biomonitoring programs. To our best knowledge, this is the first study to assemble a barcode reference library for these dominant groups of marine macroinvertebrate species from this important region of the Atlantic. However, we are still far from having a representative reference library for such diverse taxonomic groups, with prevalence of large gaps in the library. Furthermore, other important marine taxa (e.g. echinoderms or ascidians) should be included in forthcoming studies to improve the completion of reference libraries and broader integration in ecological assessments of marine species, namely through DNA metabarcoding. A significant finding emerging from our analysis was the circa 50% higher number of species delimited through molecular data (i.e. BINs) compared to described morphospecies occurring in this Atlantic Iberia only. Implications of such exceptional levels of suspected hidden diversity should be taken into consideration in upcoming macroinvertebrate-based ecosystem monitoring and research. The continuous growth of reference libraries with comprehensive sampling strategies, ranging from different regions and a broad range of specimens, combined with morphological taxonomy and molecular phylogenetic techniques will probably allow to better understand the diversity and deep genetic structure within species, in order to solve the observed discrepancies and incongruences, most of them probably associated with undescribed or cryptic diversity.

References

Abad D., Albaina A., Aguirre M., Laza-Martínez A., Uriarte I., Uriarte A., Villate F., Estonba A. (2016). Is metabarcoding suitable for estuarine plankton monitoring? A comparative study with microscopy. Marine Biology. 163(7), 149. http://doi.org/10.1007/s00227-016-2920-0.

Araújo R., Bárbara I., Tibaldo M., Berecibar E., Tapia P.D., Pereira R., Santos R., Pinto I.S. (2009). Checklist of benthic marine algae and cyanobacteria of northern Portugal. Botanica Marina. 52(1), 24-46. https://doi.org/10.1515/BOT.2009.026.

Ardura A. (2019). Species-specific markers for early detection of marine invertebrate invaders through eDNA methods: gaps and priorities in GenBank as database example. Journal for Nature Conservation. 47, 51-57. http://doi.org/10.1016/j.jnc.2018.11.005.

Aylagas E., Borja Á., Muxika I., Rodríguez-Ezpeleta N. (2018). Adapting metabarcoding based benthic biomonitoring into routine marine ecological status assessment networks. Ecological Indicators. 95, 194–202. https://doi.org/10.1016/j.ecolind.2018.07.044.

Aylagas E., Borja A., Rodríguez-Ezpeleta N. (2014). Environmental status assessment using DNA metabarcoding: towards a genetics based marine biotic index (gAMBI). PLoS One. 9(3), e90529. https://doi.org/10.1371/journal.pone.0090529.

Barco A., Raupach M.J., Laakmann S., Neumann H., Knebelsberger T. (2016). Identification of North Sea molluscs with DNA barcoding. Molecular Ecology Resources. 16(1), 288-297. https://doi.org/10.1111/1755-0998.12440.

Bergsten J., Bilton D.T., Fujisawa T., Elliott M., Monaghan M.T., Balke M., Hendrich L., Geijer J., Herrmann J., Foster G.N., Ribera I., Nilsson A.N., Barraclough T.G., Vogler, A.P. (2012). The effect of geographical scale of sampling on DNA barcoding. Systematic Biology. 61, 851–869. https://doi.org/10.1093/sysbio/sys037.

Best R.J., Stachowicz J.J. (2013). Phylogeny as a proxy for ecology in seagrass amphipods: which traits are most conserved?. PLoS One. 8(3). https://doi.org/10.1371/journal.pone.0057550.

Boaventura D., Ré P., Cancela da Fonseca L., Hawkins S.J. (2002). Intertidal rocky shore communities of the continental Portuguese coast: analysis of distribution patterns. Marine Ecology. 23(1), 69-90. https://doi.org/10.1046/j.1439-.0485.2002.02758.x.

Borges L.M.S., Hollatz C., Lobo J., Cunha A.M., Vilela A.P., Calado G., Coelho R., Costa A.C., Ferreira M.S.G., Costa M.H., Costa F.O. (2016). With a little help from DNA barcoding: investigating the diversity of Gastropoda from the Portuguese coast. Scientific Reports. 6, 20226. https://doi.org/10.1038/srep20226.

Borja A., Franco J., Pérez V. (2000). A marine biotic index to establish the ecological quality of soft-bottom benthos within European estuarine and coastal environments. Marine Pollution. Bulletin. 40, 1100–1114. https://doi.org/10.1016/S0025-326X(00)00061-8.

Bourlat S.J., Borja A., Gilbert J., Taylor M.I., Davies N., Weisberg S.B., Griffith J.F., Lettieri T., Field D., Benzie J., Glöckner F.O., Rodríguez-Ezpeleta N., Faith D.P., Bean T.P., Obst M. (2013). Genomics in marine monitoring: new opportunities for assessing marine health status. Marine Pollution Bulletin. 74(1), 19-31. http://doi.org/10.1016/j.marpolbul.2013.05.042.

Briggs J.C., Bowen B.W. (2012). A realignment of marine biogeographic provinces with particular reference to fish distributions. Journal of Biogeography. 39, 12–30. https://doi.org/10.1111/j.1365-2699.2011.02613.x.

Briski E., Ghabooli S., Bailey S.A., MacIsaac H.J. (2016). Are genetic databases sufficiently populated to detect non-indigenous species? Biological Invasions. 18, 1911-1922. http://dx.doi.org/10.1007/s10530-016-1134-1.

Cariani A., Messinetti S., Ferrari A., Arculeo M., Bonello J.J., Bonnici L., Cannas R., Carbonara P., Cau A., Charilaou C., Ouamari N.E., Fiorentino F., Follesa M.C., Garofalo G., Golani D., Guarniero I., Hanner R., Hemida F., Kada O., lo Brutto S., Mancusi C., Morey G., Schembri P.J., Serena F., Sion L., Stagioni M., Tursi A., Vrgoc N., Steinke D., Tinti F. (2017). Improving the Conservation of Mediterranean Chondrichthyans: The ELASMOMED DNA Barcode Reference Library. PLoS One. 12(1), e0170244. https://doi.org/10.1371/journal.pone.0170244.

Chainho P., Fernandes A., Amorim A., Ávila S.P., Canning-Clode J., Castro J.J., Costa A.C., Costa J.L., Cruz T., Gollasch S., Grazziotin-Soares C., Melo R., Micael J., Parente M.I., Semedo J., Silva T., Sobral D., Sousa M., Torres P., Veloso V., Costa M.J. (2015). Non-indigenous species in Portuguese coastal areas, coastal lagoons, estuaries and islands. Estuarine, Coastal and Shelf Science. 167,199–211. https://doi.org/10.1016/j.ecss.2015.06.019.

Costa F.O., deWaard J.R., Boutillier J., Ratnasingham S., Dooh R.T., Hajibabaei M., Hebert P.D.N. (2007). Biological identifications through DNA barcodes: the case of the Crustacea. Canadian Journal of Fisheries and Aquatic Sciences. 64, 272–295. https://doi.org/10.1139/f07-008.

Costa F.O., Henzler C.M., Lunt D.H., Whiteley N.M., Rock J. (2009). Probing marine *Gammarus* (Amphipoda) taxonomy with DNA barcodes. Systematics and Biodiversity. 7, 365–379. https://doi.org/10.1017/S1477200009990120.

Costa F.O., Landi M., Martins R., Costa M.H., Costa M.E., Carneiro M., Alves M.J., Steinke D., Carvalho G.R. (2012). A ranking system for reference libraries of DNA barcodes: application to marine fish species from Portugal. PLoS One. 7, e35858. https://doi.org/10.1371/journal.pone.0035858.

Cowart D.A., Pinheiro M., Mouchel O., Maguer M., Grall J., Miné J., Arnaud-Haond S. (2015). Metabarcoding is powerful yet still blind: a comparative analysis of morphological and molecular surveys of seagrass communities. PLoS One. 10(2), e0117562. https://doi.org/10.1371/journal. pone.0117562.

da Silva J.M., Dos Santos A., Cunha M.R., Costa F.O., Creer S., Carvalho G.R. (2013). Investigating the molecular systematic relationships amongst selected *Plesionika* (Decapoda: Pandalidae) from the Northeast Atlantic and Mediterranean Sea. Marine Ecology. 34, 157–170. https://doi.org/10.1111/j.1439-0485.2012.00530.x.

Danovaro R., Carugati L., Berzano M., Cahill A.E., Carvalho S., Chenuil A., Corinaldesi C., Cristina S., David R., Dell'anno A., Dzhembekova N., Garcés E., Gasol J., Goela P., Feral J.P., Ferrera I., Forster R., Kurekin A.A., Rastelli E., Marinova V., Miller P.I., Moncheva S., Newton, A., Pearman J., Pitois S., Renè A., Rodriguez-Ezpeleta N., Saggiomo V., Simis S.G.H., Stefanova K., Wilson C., Lo Martire M., Greco S., Cochrane S., Borja A. (2016). Implementing and innovating marine monitoring approaches for assessing marine environmental status. Frontiers in Marine Science. 3, 213. https://doi.org/10.3389/fmars.2016.00213.

Desiderato A., Costa F.O., Serejo C.S., Abbiati M., Queiroga H., Vieira P.E. (2019). Macaronesian islands as promoters of diversification in amphipods: The remarkable case of the family Hyalidae (Crustacea, Amphipoda). Zoologica Scripta. 48(3), 359-375. https://doi.org/10.1111/zsc.12339.

European Commission (2000). Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. https://eur-lex.europa.eu/eli/dir/2000/60/oj.

European Commission (2008). Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive). https://eur-lex.europa.eu/eli/dir/2008/56/oj.

George J.D. (1967). Cryptic polymorphism in the cirratulid polychaete, *Cirriformia tentaculata*. Journal of the Marine Biological Association of the United Kingdom. 47(1), 75-79. https://doi.org/10.1017/S0025315400033579.

Guerra-García J.M., Baeza-Rojano E., Cabezas M.P., García-Gómez J.C. (2011). Vertical distribution and seasonality of peracarid crustaceans associated with intertidal macroalgae. Journal of Sea Research. 65(2), 256-264. https://doi.org/10.1016/j.seares.2010.12.001.

Hajibabaei M., Spall J.L., Shokralla S., van Konynenburg S. (2012). Assessing biodiversity of a freshwater benthic macroinvertebrate community through non-destructive environmental barcoding of DNA from preservative ethanol. BioMed Central Ecology. 12, 28. https://doi.org/10.1186/1472-6785-12-28.

Hayward P.J., Ryland J.S. (1995). Handbook of the marine fauna of North-west Europe. Oxford University press. Oxford, United Kigdom.

Hollatz C., Leite B.R., Lobo J., Froufe H., Egas C., Costa F.O. (2017). Priming of a DNA metabarcoding approach for species identification and inventory in marine macrobenthic communities. Genome. 60, 260– 271. https://doi.org/10.1139/gen-2015-0220.

Holstein J. (2018). worms: Retriving Aphia Information from World Register of Marine Species. R package version 0.2.2. https://CRAN.R-project.org/package=worms (accessed February 2020).

Hupalo K., Teixeira M.A.L., Rewicz T., Sezgin M., Iannilli V., Karaman G.S., Grabowski M., Costa F.O. (2019). Persistence of phylogeographic footprints helps to understand cryptic diversity detected in two marine amphipods widespread in the Mediterranean basin. Molecular Phylogenetics and Evolution. 132, 53-66. https://doi.org/10.1016/j.ympev.2018.11.013.

Katoh K., Standley D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution. 30, 772–780. https://doi.org/10.1093/molbev/mst010.

Kelly R.P., Closek C.J., O'Donnell J.L., Kralj J.E., Shelton A.O., Samhouri J.F. (2017). Genetic and manual survey methods yield different and complementary views of an ecosystem Frontiers in Marine Science. 3, 283. https://doi.org/10.3389/fmars.2016.00283.

Keskin E., Atar H.H. (2013). DNA barcoding commercially important fish species of Turkey. Molecular Ecology Resources. 13, 788-797. https://doi.org/10.1111/1755-0998.12120.

Kimura M. (1980). A simple model for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution. 16, 111–120. https://doi.org/10.1007/BF01731581.

Korshunova T., Sanamyan N., Zimina O., Fletcher K., Martynov A. (2016). Two new species and a remarkable record of the genus *Dendronotus* from the North Pacific and Arctic oceans (Nudibranchia). ZooKeys. 630, 19–42. https://doi.org/10.3897/zookeys.630.10397.

Krapp-Schickel T., Vader W. (1998). What is, and what is not, *Caprella acanthifera* Leach, 1814 (Amphipoda, Caprellidae)? Part 1: the acanthifera-group. Journal of Natural History. 32:7, 949-967. http://dx.doi.org/10.1080/00222939800770491.

Langeneck J., Scarpa F., Maltagliati F., Sanna D., Barbieri M., Cossu P., Mikac B., Galletti M.C., Castelli A., Casu M. (2020). A complex species complex: The controversial role of ecology and biogeography in the evolutionary history of *Syllis gracilis* Grube, 1840 (Annelida, Syllidae). Journal of Zoological Systematics and Evolutionary Research. 58(1), 66-78. https://doi.org/10.1111/jzs.12336.

Layton K.K., Martel A.L., Hebert P.D.N. (2014). Patterns of DNA barcode variation in Canadian marine molluscs. PloS One. 9(4), e95003. https://doi.org/10.1371/journal.pone.0095003.

Leese F., Bouchez A., Abarenkov K., Altermatt F., Borja A., Bruce K., Torbjørn E., Ciampor F., Ciamporová-Zaťovicová Z., Costa F.O., Duarte S., Elbrecht V., Fontaneto D., Franc A., Geiger M.F., Hering D., Kahlert M., Stroil B.K., Kelly M., Keskin E., Liska I., Mergen P., Meissner K., Pawlowski J., Penev L., Reyjol Y., Rotter A., Steinke D., van der Wal B., Vitecek S., Zimmermann J., Weigand A.M. (2018). Why we need sustainable networks bridging countries, disciplines, cultures and generations for aquatic biomonitoring 2.0: a perspective derived from the DNAqua-Net COST action. Advances in Ecological Research. 58, 63-99. https://doi.org/10.1016/bs.aecr.2018.01.001.

Lejzerowicz F., Esling P., Pillet L., Wilding T.A. (2015). High-throughput sequencing and morphology perform equally well for benthic monitoring of marine ecosystems. Scientific Reports. 5, 13932. https://doi.org/10.1038/srep13932.

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. 112(7), 2076-2081. https://doi.org/10.1073/pnas.1424997112.

Leray M., Knowlton N. (2016). Censusing marine eukaryotic diversity in the twenty-first century. Philosophical Transactions of the Royal Society B. 371, 20150331. http://doi.org/10.1098/rstb.2015.0331.

Lima F.P., Ribeiro P.A., Queiroz N., Hawkins S.J., Santos A.M. (2007). Do distributional shifts of northern and southern species of algae match the warming pattern?. Global Change Biology. 13(12), 2592-2604. https://doi.org/10.1111/j.1365-2486.2007.01451.x.

Lincoln R. (1979). British marine amphipoda: Gammaridea. British Museum (Natural History). London, United Kigdom.

Lobo J., Ferreira M.S.G., Antunes I.C., Teixeira M.A.L., Borges L.M., Sousa R., Gomes P.A., Costa M.H., Cunha M.R., Costa F.O. (2017). Contrasting morphological and DNA barcode-suggested species boundaries among shallow-water amphipod fauna from the southern European Atlantic coast. Genome. 60(2), 147-157. https://doi.org/10.1139/gen-2016-0009.

Lobo J., Teixeira M.A.L., Borges L.M., Ferreira M.S.G., Hollatz C., Gomes P.T., Sousa R., Ravara A., Costa M.H., Costa F.O. (2016). Starting a DNA barcode reference library for shallow water polychaetes from the southern European Atlantic coast. Molecular Ecology Resources. 16(1), 298-313. https://doi.org/10.1111/1755-0998.12441.

McCarthy J.B., Krug P.J., Valdés Á. (2019). Integrative systematics of *Placida cremoniana* (Trinchese, 1892) (Gastropoda, Heterobranchia, Sacoglossa) reveals multiple pseudocryptic species. Marine Biodiversity. 49(1), 357-371. https://doi.org/10.1007/s12526-017-0812-2.

Miralles L., Ardura A., Arias A., Borrell Y.J., Clusa L., Dopico E., Hernandez de Rojas A., Lopez B., Muñoz-Colmenero M., Roca A., Valiente A.G., Zaiko A., Garcia-Vazquez E. (2016). Barcodes of marine invertebrates from north Iberian ports: Native diversity and resistance to biological invasions. Marine Pollution Bulletin. 112(1-2), 183-188. http://doi.org/10.1016/j.marpolbul.2016.08.022.

Morrow C.C., Thorpe J.P., Picton B. (1992). Genetic divergence and cryptic speciation in two morphs of the common subtidal nudibranch *Doto coronata* (Opisthobranchia: Dendronotacea: Dotoidae) from the northern Irish Sea. Marine Ecology Progress Series. 84, 53–61. http://doi.org/10.3354/meps084053.

Múrria, C., Väisänen L.O., Somma S., Wangensteen O.S., Arnedo M.A., Prat N. (2019). Towards an Iberian DNA barcode reference library of freshwater macroinvertebrates and fishes. Limnetica. 39(1), 73-92. http://doi.org/10.23818/limn.39.06.

Naylor E. (1972). British Marine Isopods. Keys and Notes for the Identification of the Species– Synopses of the British Fauna. Academc Press. London-New York.

Nunes A.L., Katsanevakis S., Zenetos A., Cardoso A.C. (2014). Gateways to alien invasions in the European seas. Aquatic Invasions. 9:133–44. http://doi.org/10.3391/ai.2014.9.2.02.

Oliveira L.M., Knebelsberger T., Landi M., Soares P., Raupach M.J., Costa F.O. (2016). Assembling and auditing a comprehensive DNA barcode reference library for European marine fishes. Journal of fish biology. 89(6), 2741-2754. https://doi.org/10.1111/jfb.13169.

Pascual S., Villalba A., Abollo E., Garci M., González A.F., Nombela M., Posada D., Guerra A. (2010). The mussel *Xenostrobus securis*: a well-established alien invader in the Ria de Vigo (Spain, NE Atlantic). Biological Invasions. 12(7), 2091–103. http://doi.org/10.1007/s10530-009-9611-4.

Pawlowski J., Kelly-Quinn M., Altermatt F., Apothéloz-Perret-Gentil L., Beja P., Boggero A., Borja A., Bouchez A., Cordier T., Domaizon I., Feio M.J., Filipe A.F., Fornaroli R., Graf W., Herder J., van der Hoorn B., Jones J.I., Sagova-Mareckova M., Moritz C., Barquín J., Piggott J.J., Pinna M., Rimet F., Rinkevich B., Sousa-Santos C., Specchia V., Trobajo R., Vasselon V., Vitecek S., Zimmerman J., Weigand A.M., Leese F., Kahlert M. (2018). The future of biotic indices in the ecogenomic era: Integrating (e)DNA metabarcoding in biological assessment of aquatic ecosystems. Science of the Total Environment. 637, 1295-1310. https://doi.org/10.1016/j.scitotenv.2018.05.002.

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

Pejovic I., Ardura A., Miralles L., Arias A., Borrell Y.J., Garcia-Vazquez E. (2016). DNA barcoding for assessment of exotic molluscs associated with maritime ports in northern Iberia. Marine Biology Research. 12:2, 168-176. https://doi.org/10.1080/17451000.2015.1112016.

Pereira S.G., Lima F.P., Queiroz N.C., Ribeiro P.A., Santos A.M. (2006). Biogeographic patterns of intertidal macroinvertebrates and their association with macroalgae distribution along the Portuguese coast. Hydrobiologia. 555(1), 185. https://doi.org/10.1007/s10750-005-1115-3.

Pola M., Gosliner T.M. (2010). The first molecular phylogeny of cladobranchian opisthobranchs (Mollusca, Gastropoda, Nudibranchia). Molecular Phylogenetics and Evolution. 56, 931–941. http://dx.doi.org/10.1016/j.ympev.2010.05.003.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/ (accessed February 2020)

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

Ratnasingham S., Hebert P.D.N. (2013). DNA-based registry for all animal species: the Barcode Index Number (BIN) system. PLoS One. 8, e66213. https://doi.org/10.1371/journal.pone.0066213.

Raupach M.J., Barco A., Steinke D., Beermann J., Laakmann S., Mohrbeck I., Neumann H., Kihara T.C., Pointner K., Radulovici A., Segelken-Voigt A., Wesse C., Knebelsberger T. (2015). The application of DNA barcodes for the identification of marine crustaceans from the North Sea and adjacent regions. PLoS One. 10(9): e0139421. https://doi.org/10.1371/journal.pone.0139421.

Rey A., Basurko O.C., Rodriguez-Ezpeleta N. (2019). Guidelines and considerations for metabarcoding-based port baseline biodiversity surveys: towards improved marine non-indigenous species monitoring. bioRxiv. http://doi.org/10.1101/689307.

Rubal M., Costa-Garcia R., Besteiro C., Sousa-Pinto I., Veiga P. (2018). Mollusc diversity associated with the non-indigenous macroalga *Asparagopsis armata* Harvey, 1855 along the Atlantic coast of the Iberian Peninsula. Marine Environmental Research. 136, 1-7. https://doi.org/10.1016/j.marenvres.2018.02.025.

Ruffo S. (1982). The Amphipoda of the Mediterranean. Memoires de l'Institut Oceanographique. Monaco.

Scott C. (2019). bold: Interface to Bold Systems API. R package version 0.9.0. https://CRAN.R-project.org/package=bol (accessed February 2020).

Shipman C., Gosliner T. (2015). Molecular and morphological systematics of *Doto* Oken, 1851 (Gastropoda: Heterobranchia), with descriptions of five new species and a new genus. Zootaxa. 3973(1), 57-101. http://dx.doi.org/10.11646/zootaxa.0000.0.0.

Siddall M.E., Fontanella F.M., Watson S.C., Kvist S., Erséus C. (2009). Barcoding bamboozled by bacteria: convergence to metazoan mitochondrial primer targets by marine microbes. Systematic Biology. 58(4), 445-451. https://doi.org/10.1093/sysbio/syp033.

Spalding M.D., Fox H.E., Gerald G.R., Davidson N., Ferdaña Z.A., Finlayson M., Halper B.S., Jorge M.A., Lombana A.I., Lourie S.A., Martin K.D., McManus E., Molnar J., Recchia C.A., Robertson J. (2007). Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. BioScience. 57, 573–583. https://doi.org/10.1641/B570707.

Tang C.Q., Leasi F., Obertegger U., Kieneke A., Barraclough T.G., 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. 109(40), 16208-16212. https://doi.org/10.1073/pnas.1209160109.

Teixeira M.A.L., Vieira P.E., Pleijel F. Sampieri B.R., Ravara A., Costa F.O., Nygren A. (2019). Molecular and morphometric analyses identify new lineages within a large *Eumida* (Annelida) species complex. Zoologica Scripta. 49, 222-235. https://doi.org/10.1111/zsc.12397.

Trickey J.S., Thiel M., Waters J.M. (2016). Transoceanic dispersal and cryptic diversity in a cosmopolitan rafting nudibranch. Invertebrate Systematics. 30(3), 290-301. https://doi.org/10.1071/IS15052.

Veiga P., Torres A.C., Aneiros F., Sousa-Pinto I., Troncoso J.S., Rubal M. (2016). Consistent patterns of variation in macrobenthic assemblages and environmental variables over multiple spatial scales using taxonomic and functional approaches. Marine Environmental Research. 120, 191-201. https://doi.org/10.1016/j.marenvres.2016.08.011.

Viard F., Roby C., Turon X., Bouchemousse S., Bishop J. (2019). Cryptic diversity and database errors challenge non-indigenous species surveys: an illustration with *Botrylloides* spp. in the English Channel and Mediterranean Sea. Frontiers in Marine Science. 6, 615. https://doi.org/10.3389/fmars.2019.00615.

Vieira P.E., Desiderato A., Holdich D.M., Soares P., Creer S., Carvalho G.R., Costa F.O., Queiroga H. (2019). Deep segregation in the open ocean: Macaronesia as an evolutionary hotspot for low dispersal marine invertebrates. Molecular Ecology. 28(7), 1784-1800. https://doi.org/10.1111/mec.15052.

Weigand H., Beermann A.J., Ciampor F., Costa F.O., Csabai Z., Duarte S., Geiger M.F., Grabowski M., Rimet F., Rulik B., Strand M., Szucsich N., Weigand A.M., Willassen E., Wyler S.A., Bouchez A., Borja A., Ciamporová-Zaťovicová Z., Ferreira S., Dijkstra K.B., Eisendle U., Freyhof J., Gadawski P., Graf W., Haegerbaeumer A., van der Hoorn B.B., Japoshvili B., Keresztes L., Keskin E., Lesse F., Macher J.N., Mamos T., Paz G., Pesic V., Pfannkuchen D.M., Pfannkuchen M.A., Price B.W., Rinkevich B., Teixeira M.A.L., Várbíró G., Ekrem T. (2019). DNA barcode reference libraries for the monitoring of aquatic biota in Europe: Gap-analysis and recommendations for future work. Science of the Total Environment. 678, 499-524. https://doi.org/10.1016/j.scitotenv.2019.04.247.

Chapter 3

Comparing species detection success between molecular markers and primers in DNA metabarcoding of coastal macroinvertebrate communities

Abstract

DNA metabarcoding has been shown to have great potential to improve marine biomonitoring programs by providing a rapid and accurate assessment of species composition in zoobenthic communities. However, some methodological improvements are still required, especially related to failed detections and primers efficiency, taxonomic discrimination ability and incompleteness of databases. Here we assessed the efficiency of two different marker loci (COI and 18S) and four different primer-pairs (COI only) in marine species detection through DNA metabarcoding of the macroinvertebrate communities colonizing three types of artificial substrates (slate, PVC and granite), deployed in Toralla Island, NW Spain. To accurately compare detection success between COI and 18S we also verified if reference sequences of the species detected were present in each marker's database, respectively BOLD and SILVA. Globally, a higher number of species was detected with COI than 18S (104 vs 90), however, the single 18S primer detected more species than each of the COI primers individually (91 species for 18S vs 84 and 64 species for the two COI primers). We recorded extensive complementarity in the species detected by each marker (ranging only from 9 to 13 overlapping detections in the different substrates), with 70.1% of the species detected exclusively by 18S or COI. Most of the detected species have reference sequences in their respective databases (81.3% for COI and 74.3% for 18S), meaning that when a species was detected by one marker and not by the other it was most likely due to faulty amplification, and not by lack of matching sequences in the database. However, we identified 10 cases of detected species which were not present in both databases. Overall, results showed the impact of marker and primer applied on species detection ability and indicated that, currently, no single marker is able to fully detect the diversity of marine zoobenthic communities. Although detection success can be improved through broader species representation in the reference libraries, our results indicated that primer amplification bias will still impede full diversity diagnosis when using a single marker.

Keywords: DNA metabarcoding; COI and 18S genes; Primer efficiency; Marine macroinvertebrate diversity; Database completeness.

3.1 Introduction

DNA metabarcoding is the identification of a species present in a bulk sample through the use of DNA barcoding (i.e. DNA amplification of standard regions of a genome) coupled with high-throughput sequencing (HTS) (Taberlet et al., 2012). DNA metabarcoding studies have been developed for diverse taxonomic groups (e.g.: terrestrial arthropods: Elbrecht et al., 2019; freshwater macroinvertebrates: Bista et al., 2018; Giebner et al., 2020; meiofaunal organisms: Fais et al., 2020b; marine communities: Leray and Knowlton, 2015; Aylagas et al., 2018; Ip et al., 2019; Chang et al., 2020a), using a wide range of laboratory procedures (Andújar et al., 2018) and addressing questions about species richness, taxonomic composition, as well as biodiversity patterns (McGee et al., 2019; Piñol et al., 2019).

Metabarcoding allows for comparison across studies, however the harmonization and standardization of protocols is still far from being established (van der Loos and Nijland, 2020). While DNA-based approaches for assessing and monitoring marine macroinvertebrate species are constantly evolving (Andújar et al., 2018), the diversity of the adopted methodologies, including the use of different primer-pairs or molecular markers, the lack of accurate and complete reference databases, and the continuous emergence of new sequencing innovations and bioinformatics pipelines implies low standardization and comparability among studies (Cowart et al., 2015; Leray and Knowlton, 2017), which remain important challenges that should be addressed.

Targeting marine species is specially challenging due to the broad taxonomic and phylogenetic composition of marine communities, and the choice of marker usually depends on the target taxa. However, the balance between the range of taxonomic coverage and the taxonomic discrimination ability should be considered in the choice of target genomic region and/or primer-pairs, since it may affect the number of species and the taxonomic groups detected, as well as the accuracy of species identification (van der Loos and Nijland, 2020). Furthermore, a high proportion of taxonomically unassigned reads have been reported in a number of studies (e.g. Aylagas et al., 2018; Wangensteen et al., 2018), and uncertainty remain as to whether this is related to unknown diversity, PCR biases, sequencing errors, presence of pseudogenes, unspecific amplification or incompleteness of databases (Derycke et al., 2020).

Initially, a fundamental and critical decision in a DNA metabarcoding study is centered on which genomic region should be targeted. The 5' end of the mitochondrial COI gene (COI-5P) is the standard barcode for animal life and the backbone of the universal Barcode of Life species identification system (Hebert et al., 2003), being the recommend marker for community metabarcoding (Andújar et al., 2018) and by far the

most well represented genomic region in public databases (Porter and Hajibabaei, 2020). After COI-5P, the nuclear small subunit rRNA gene (18S) is among the most widely used markers in marine biodiversity studies (e.g. Lejzerowicz et al., 2015; Wangensteen et al., 2018; Zhang et al., 2018; Fais et al., 2020a).

PCR-based methodologies are highly influenced by amplification biases thereby encouraging the use of several primer-pairs in different metabarcoding studies (Bista et al., 2018; Elbrecht et al., 2019; Hajibabaei et al., 2019; Porter and Hajibabaei, 2020). Primer design can be challenging, due to the trade-off between taxonomic scope (i.e. able to amplify all the species present on the sample) and specificity (i.e. able to amplify specific target species and rare species). Amplicon length, primer mismatches, GC content, and polymerase errors can also affect the ability to detect the species present on a marine sample (Kebschull and Zador, 2015; Piñol et al., 2015; Derycke et al., 2020). Due to its high variation, primer design is particularly challenging in the case of the COI-5P barcode region (Deagle et al., 2014), prompting the search for alternative regions that match highly conserved binding sites with a presumed more even success across taxa (e.g. 18S rRNA gene).

Perhaps the most common widely used broad-range primers (i.e. potential to amplify a DNA fragment across a broad taxonomic scope) are the ones designed by Folmer et al. (1994) - LCO1490/HCO2198. However, even broad-range primers demonstrated more affinity for some species and consequently do not perfectly match the DNA of all species present in a bulk sample. Low species detection ability has been associated with the use of nondegenerate primers (Clarke et al., 2014; Elbrecht et al., 2017; Collins et al., 2019). Numerous alternative pairs of degenerate primers have been proposed for the entire 658 bp of the COI-5P barcode (e.g. jgLCO1490/jgHCO2198 - Geller et al., 2013; LoboF1/LoboR1 - Lobo et al., 2013). The Leray-Geller fragment (mlCOIintF/jgHCO2198 - Leray et al., 2013) is a degenerate primer-pair widely used in DNA metabarcoding studies for different type of taxa (Leray and Knowlton, 2015; Clarke et al., 2017; Ransome et al., 2017; Aylagas et al., 2018), mostly due to their design for marine organisms with a wide phylogenetic coverage and fair amplicon length (313 bp). The combination of mlCOIintF with LoboR1 amplifies exactly the same fragment and with demonstrated success in the amplification of DNA barcodes of marine taxa (Hollatz et al., 2017; Chang et al., 2020a). The combination of non- with degenerate primers (LCO1490/III_C_R; 325 bp) was also successfully tested for a different variety of taxa (Zhang et al., 2018; Elbrecht et al., 2019; Piñol et al., 2019), including marine species (Lacoursière-Roussel et al., 2019). Primers amplifying the nuclear genes constitute alternatives to consider, mainly because of their slower rate of evolution which results in more conserved regions and facilitates the design

of primers. However, reference databases for such genomic regions are less populated (Andújar et al., 2018) and COI-primers often outperformed primers for rDNA loci on taxon recovery (Clarke et al., 2017; Elbrecht et al., 2017, 2019) and species discrimination ability (Tang et al., 2012; Clarke et al., 2017).

For DNA metabarcoding studies, multiple sets of primers amplifying different molecular markers have been used to target a broad of taxonomic groups in different marine communities (Dowle et al., 2015; Leray and Knowlton, 2015; Zaiko et al., 2015; Ip et al., 2019; Giebner et al., 2020). However, the majority of studies use a single primer-pair or single marker loci strategy (Aylagas et al., 2018; Chang et al., 2020a). In a study which comprehensively reviews DNA metabarcoding studies over the last 10 years, the authors concluded that only 25% of the publications used more than one marker (van der Loos and Nijland, 2020). Problems associated to non-detection of specific taxonomic groups could be related to taxon-specific problems (e.g. difficulties in DNA extraction of mollusks; van der Loos and Nijland, 2020). Furthermore, for marine macroinvertebrate communities', which are typically among the most phylogenetically diverse, the use of multiple primer-pairs or a multi marker approach is advocated to improve species assignment and community inventories (Cowart et al., 2015; Drummond, 2015; Alberdi et al., 2017; Hollatz et al., 2017; Wangensteen et al., 2018).

Although DNA metabarcoding studies aim species-level assignments (Taberlet et al., 2012), the existence of gaps in the reference sequence databases (Weigand et al., 2019), associated with the lack of species-level discrimination for some markers (e.g. 18S rRNA; von Ammon et al., 2018), reduces the resolution level of taxonomic identifications and lead to identifications at higher taxonomic ranks (Porter and Hajibabaei, 2020). The taxonomic coverage and quality of available reference sequences for marine macroinvertebrate species vary according to marker and database. For DNA-based study of marine macroinvertebrate communities based on COI-5P barcode region, Barcode of Life Data system (BOLD - http://v4.boldsystems.org/index.php; Ratnasingham and Hebert, 2007) is the prime database, constituting the most populated with reference sequences (Alberdi et al., 2017; Derycke et al., 2020). For 18S rDNA sequences, one of the most prominent databases is SILVAngs (https://ngs.arb-silva.de/silvangs/, Quast et al., 2013) which consists on a curated public database, based on sequences from small and large subunit of eukaryotes (as well as bacteria and archaea), and the taxonomic assignment is based on phylogenetic placement (Pruesse et al., 2007).

Considering the importance of choice of marker and primer to improve taxonomic coverage and resolution of DNA metabarcoding, we investigated the impact of these factors on the composition and structure of

marine macroinvertebrate communities. We first evaluated the performance of four primer-pairs, targeting internal segments within the COI-5P region, on marine macroinvertebrate detection, to select the best strategy to characterize these communities. From these analyses, we selected three different primer-pairs targeting COI-5P and 18S loci, to compare their ability to detect macroinvertebrates at species-level and to evaluate the benefits of the use of two molecular markers on species recovery rate. Finally, we also conducted an assessment of the availability of reference sequences for all species detected in the study, in order to identify the existence of gaps in both databases (BOLD for COI-5P and SILVA for 18S rRNA gene) and attempting to infer the reasons for failed detections.

3.2 Materials and methods

3.2.1 Sampling design

This study was developed in Ría de Vigo, a semi-enclosed heavily populated bay on the NW coast of Spain. This area includes both hard and soft substrata, which have a high primary productivity due to the influence of coastal upwelling-downwelling dynamics (Prego and Fraga, 1992). Affected by several human activities (e.g. sewage runoff or harvesting) and constituted by important busy harbours, Ría de Vigo shows a funnel-like morphology (SW-NE direction) where Cíes Islands acts as a shelter against waves (Veiga et al., 2016).

In December 2016, four replicates (flat panels 10 x 10 cm) of three different types of artificial substrates - slate, polyvinyl chloride (PVC) and granite - were randomly deployed on the dock of Toralla Island (42° 12' 2.267'', 8° 48' 4.187''), approximately 1.5 m below water surface. Using a hermetic plastic bag, after 3, 7, 10 and 15 months one replicate of each substrate was randomly removed. At the laboratory, the samples were individually photographed, and the representative mobile and sessile fauna were separated; while the mobile fauna was sieved using a 500 μm mesh, the sessile fauna were scraped with a spatula into a tray. The samples were then preserved in ethanol and stored at -20 ^{0}C until further analysis.

3.2.2 DNA metabarcoding-based taxonomic identification

3.2.2.1 DNA extraction, PCR amplification and HTS procedures

DNA extraction procedures were adapted from Ivanova et al. (2006). To extract the DNA from each sample, the ethanol was filtered. Then, based on the wet weight of each sample (Braukmann et al., 2019) the appropriate volume of lysis buffer solution (100 mM NaCL, 50 mM Tris-HCL pH 8.0, 10 mM EDTA, 0.5% SDS) were added and the samples were incubated at 56 °C overnight at 200 rpm. To maximize diversity recovery, two-aliquots of each lysate was used, totalling two DNA extractions per sample. After extraction the aliquots of genomic DNA for the same sample were pooled in a single microtube and sent for high-througput sequencing (HTS).

The production of amplicon libraries and the high-throughput sequencing (HTS) were carried out at Genoinseq (Cantanhede, Portugal), as described below. First, a preliminary assessment of primer amplification efficiency of cytochrome c oxidase I was conducted to test multiple COI-5P primer-pairs that have been previously used in DNA metabarcoding studies (Table 3.1). A subset of six replicates of the substrates were selected: three replicates of slate substrate and other three replicates of PVC, all collected after three, seven and 15 months of deployment. For the primer-pair without inosines (mlCOIintF/LoboR1), PCR reactions were performed for each sample using KAPA HIFI HotStart PCR Kit according to manufacturer instructions, 0.3 μM of each PCR primer and 50 ng of template DNA in a total volume of 25 μL. For the other three primers, PCR reactions were performed using 1x Advantage 2 Polymerase Mix (Clontech, Mountain View, CA, USA), 0.2 μM of each PCR primer and 50 ng of template DNA for LoboF1/230R and mlCOIintF/jgHCO2198 and 25 ng for LCO1490/III_C_R, in a total volume of 25 μL. Negative PCR controls were included for all amplification procedures.

Second PCR reactions added indexes and sequencing adapters to both ends of the amplified target region according to manufacturer's recommendations (Illumina, 2013). PCR products were then one-step purified and normalized, pooled and pair-end sequenced in an Illumina MiSeq® sequencer with the V3 chemistry, according to manufacturer's instructions (Illumina, San Diego, CA, USA).

Table 3.1. Primer-pairs used to test the efficiency of COI-5P to amplify and assess marine macroinvertebrate species. F – forward; R – reverse; bp – base pairs.

Primer combinations and length	Direction (5'-3')	Reference	PCR thermal cycling conditions*
LCO1490 lll_C_R (325 bp)	(F) GGTCAACAAATCATAAAGATATTGG (R) GGIGGRTAIACIGTTCAICC	Folmer et al., 1994 Shokralla et al., 2015	(1) 94 °C (5 min); (2) 35 cycles: 94 °C (30 s), 52 °C (90 s), 68 °C (60 s); (3) 68 °C (10 min).
LoboF1 230R (230 bp)	(F) KBTCHACAAAYCAYAARGAYATHGG (R) CTTATRTTRTTTATICGIGGRAAIGC	Lobo et al., 2013 Gibson et al., 2015	(1) 94 °C (5 min); (2) 35 cycles: 94 °C (30 s), 48 °C (90 s), 68 °C (60 s); (3) 68 °C (10 min).
mlCOIintF LoboR1 (313 bp)	(F) GGWACWGGWTGAACWGTWTAYCCYCC (R) TAAACYTCWGGRTGWCCRAARAAYCA	Leray et al., 2013 Lobo et al., 2013	(1) 95 °C (3 min); (2) 35 cycles: 98 °C (20 s), 60 °C (30 s), 72 °C (30 s); (3) 72 °C (5 min).
mlCOIintF jgHCO2198 (313 bp)	(F) GGWACWGGWTGAACWGTWTAYCCYCC (R) TAIACYTCIGGRTGICCRAARAAYCA	Leray et al., 2013 Geller et al., 2013	(1) 94 °C (5 min); (2) 35 cycles: 94 °C (30 s), 58 °C (90 s), 68 °C (60 s); (3) 68 °C (10 min).

*Used in this study.

After the first screening of primer amplification success, two primer-pairs targeting the COI-5P barcode region and one primer-pair targeting the 18S rRNA gene were selected to amplify the marine macroinvertebrate communities from each sample (Table 3.2). PCR reactions for COI primer-pairs were the same previously described. For the 18S V4 region, PCR reactions were performed for each sample using KAPA HIFI HotStart PCR Kit according to manufacturer instructions, 0.3 µM of each PCR primer and 12.5 ng of template DNA in a total volume of 25 µL. Negative and positive PCR controls were included for all amplification procedures.

Second PCR reactions added indexes and sequencing adapters to both ends of the amplified target region according to manufacturer's recommendations (Illumina, 2013). PCR products were then one-step purified and normalized, pooled and pair-end sequenced in an Illumina MiSeq® sequencer with the V3 chemistry, according to manufacturer's instructions (Illumina, San Diego, CA, USA).

The granite sample of mobile fauna from 3 months of deployment did not amplify with mlCOIintF/LoboR1 primer-pair and were not included for further analysis.

Table 3.2. Primer-pairs and respective thermal cycling conditions used in this study to amplify marine macroinvertebrate communities. F – forward; R – reverse; bp – base pairs.

Primer combinations and length		Direction (5'-3')	Reference	PCR thermal cycling conditions
COI	LCO1490 Ill_C_R (325 bp)	(F) GGTCAACAAATCATAAAGATATTGG (R) GGIGGRTAIACIGTTCAICC	Folmer et al., 1994 Shokralla et al., 2015	(1) 94 °C (5 min); (2) 35 cycles: 94 °C (30 s), 52 °C (90 s), 68 °C (60 s); (3) 68 °C (10 min).
	mlCOIintF LoboR1 (313 bp)	(F) GGWACWGGWTGAACWGTWTAYCCYCC (R) TAAACYTCWGGRTGWCCRAARAAYCA	Leray et al., 2013 Lobo et al., 2013	(1) 95 °C (3 min); (2) 35 cycles: 98 °C (20 s), 60 °C (30 s), 72 °C (30 s); (3) 72 °C (5 min).
18S	TAReuk454FWD1 TAReukREV3 (400 bp)	(F) CCAGCASCYGCGGTAATTCC (R) ACTTTCGTTCTTGATYRA	Stoeck et al., 2010; Lejzerowicz et al., 2015	(1) 95 °C (3 min); (2) 10 cycles: 98 °C (20 s), 57 °C (30 s), 72 °C (30 s); (3) 25 cycles: 98 °C (20 s), 47 °C (30 s), 72 °C (30 s); (4) 72 °C (5 min).

3.2.2.2 Data processing

DNA sequences were quality and size filtered to remove sequencing adapters (PRINSEQ v.0.20.4, Schmieder and Edwards, 2011) and primers, to determine a minimum sequence length (150 base pairs) and to reduce sequencing biases and PCR errors (mothur v.1.39.5, Schloss et al., 2009; Kozich et al., 2013). The resultant forward and reverse reads were merged by overlapping pair-end reads with AdapterRemoval v.2.1.5 (Schubert et al., 2016). The usable reads were then processed in two pipelines of public databases: a) COI reads were submitted to mBrave – Multiplex Barcode Research and Visualization Environment (www.mbrave.net, Ratnasingham, 2019), which is based on BOLD (Ratnasingham and Hebert, 2007); b) 18S reads were analyzed by SILVAngs database (https://ngs.arb-silva.de/silvangs/, Quast et al., 2013). In both databases, taxonomic assignments were attributed when displaying ≥97% similarity with reference sequences. Only reads with match at species-level were used for further analysis, and singletons and rare sequences (i.e. less than 8 sequences) were discarded (Lobo et al., 2017). Any read that matched to non-metazoan was also excluded. The validity of the species names was verified in World Register of Marine Species (WoRMS) database.

3.2.3 Community analyses

The proportion of overlapping and unique species-level identification was determined for each primer-pair between substrates, sampling times and both using venn diagrams (http://www.venndiagrams.net/).

Multivariate analyses were carried out considering presence/absence of the taxa due to the qualitative nature of the molecular data. A two-way ANOVA was performed to assess the effect of primer-pair and marker loci on the number of marine macroinvertebrate species recovered (GraphPad Software, Inc.). Bray-Curtis measure of similarity for presence/absence of species was used to compare species identifications between substrates, and to investigate differences between primer-pairs and among the four sampling times (PRIMER v6.1.1.16; Primer-E Ltd, Plymouth, UK; Clarke and Gorley, 2006). The hierarchical clustering (CLUSTER; linkage method: UPGMA) was performed to investigate the marine macroinvertebrate community structure for each primer-pair, between substrates and sampling times (PRIMER v6.1.1.16, Primer-E Ltd, Plymouth, UK). A multivariate ANOVA based on similarities (PERMANOVA v1.0.6; Primer-E Ltd, Plymouth, UK) was tested (method: Bray-Curtis, number of permutations: 1000), to obtain the effects of primer-pairs on macroinvertebrate community's structure.

3.2.4 Gap-analysis and species discrimination ability

To perform a gap-analysis of the barcoded species, we compared the species detected on the present study with publicly available sequence records on both databases. All the available COI-5P sequences matching the detected species names were mined (on 19 October 2020) from the BOLD (Ratnasingham and Hebert, 2007) using the R package ''bold'' (Chamberlain, 2019). To assess which species have representative sequences in the SILVA database (Quast et al., 2013), all the Animalia records were mined directly from the database (on 19 October 2020). A species was considered represented if at least one sequence was available.

Although the 97% threshold is usually accepted and used as the optimal threshold in the taxonomic assignment when using COI (Flynn et al., 2015), no optimal threshold exists for 18S. Nuclear ribosomal markers usually have lower mutation rates than mitochondrial ones (e.g. Vieira et al., 2019), which suggests that a higher threshold should be used. However, if a higher threshold is used, sequences could be inaccurately attributed to a species-level identification. In marine invertebrates, a threshold between 97 and 99% is usually applied and adequate to identify and discriminate species (e.g. Brown et al., 2015; Duarte et al., 2021). To have a higher species-level taxonomic assignment success, we decided to apply the minimum value (97%). However, to confirm that a 97% threshold was adequate to obtain certainty of the identifications at the species-level and was adequate to distinguish congeneric species (i.e. species belonging to the same genus), we assess the genetic distances (Kimura 2-parameter, 1000 bootstraps), calculated in MEGA v7.0 (Kumar et al., 2016), between the aligned sequences (Clustal W; Thompson et al., 1994) of all the species belonging to the assigned genus.

3.3 Results

3.3.1 COI-5P primer-pairs testing

The first screening of primer amplification success tested four primer-pairs targeting the COI-5P region in six samples of benthic communities, and was able to identify a total of 111 different species, belonging to 9 different animal phyla: Annelida, Bryozoa, Crustacea, Echinodermata, Hydrozoa, Mollusca, Nemertea, Porifera and Tunicata. The highest amplification success was observed for the primer-pair mlCOIintF/LoboR1 (67.6% of the species), followed by mlCOIintF/jgHCO2198 (54.1% of the species), and then LCO1490/III_C_R (50.5% of the species). The lowest amplification success was recorded for the LoboF1/250R primer-pair, corresponding to 49.5% of the detected species (Fig. 3.1). The four primers displayed some complementarity in their ability to detect different taxonomic groups (Fig. 3.2.A). For example, Porifera was only recovered by mlCOIintF/LoboR1, whereas Tunicata was only detected by LoboF1/250R.

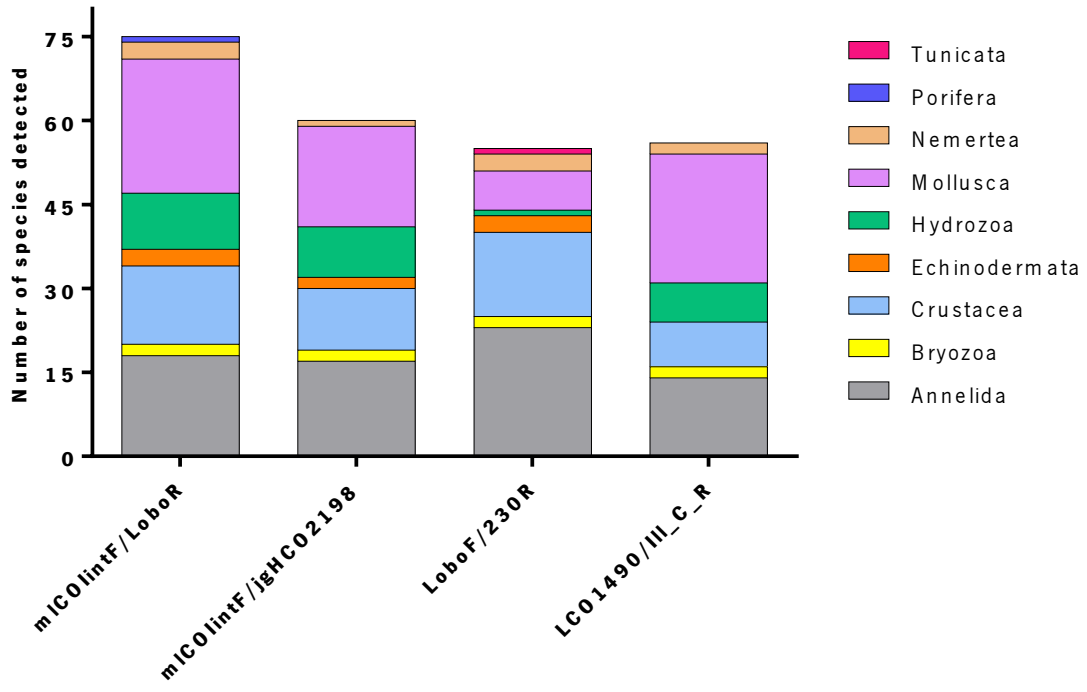


Figure 3.1. Number of marine macroinvertebrate species detected in each substrate and sampling time by each of the four primer-pairs used in the first screening of primer performance.

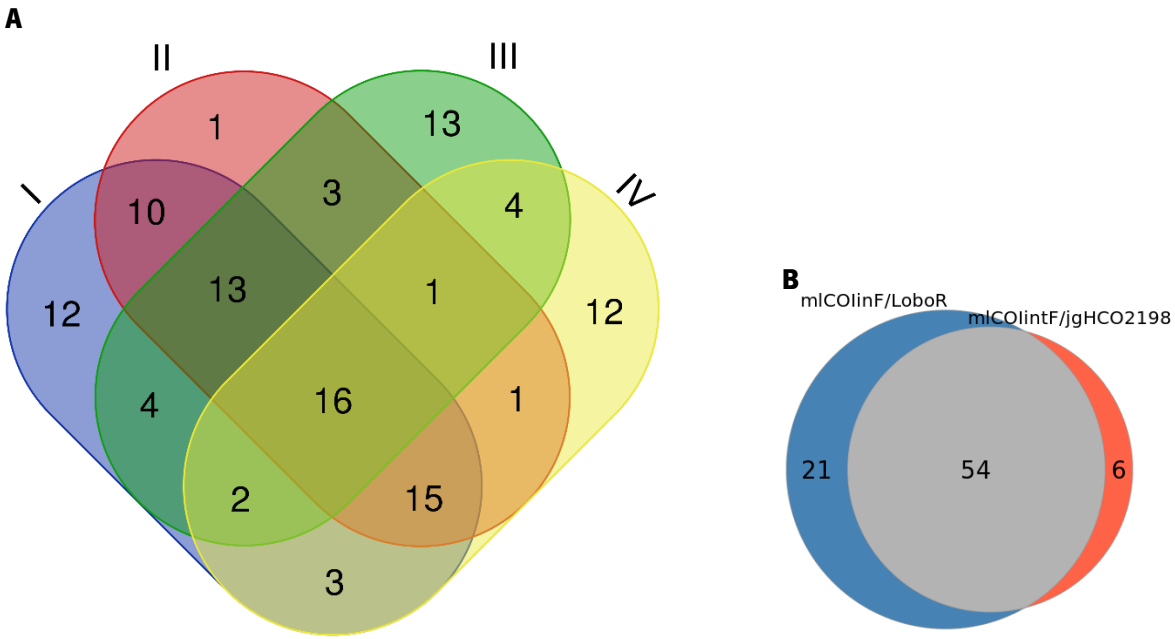


Figure 3.2. Shared and unique marine macroinvertebrate species detected (A) by the four COI-5P primer-pairs (I – mlCOIintF/LoboR1; II – mlCOIintF/jgHCO2198; III – LCO1490/III_C_R; IV – LoboF1/230R), and (B) by the two primer-pairs targeting the 3' end (mlCOIintF/LoboR1 and mlCOIintF/jgHCO2198).

3.3.2 Effect of marker and primer choice on species detection

High-throughput sequencing from marine macroinvertebrate samples, for both markers and three primer-pairs, generated a total of 2,356,818 usable reads. Of these, 53.6% were assigned at species-level to marine macroinvertebrate species: 48.5% using mlCOIintF/LoboR1, 38% with LCO1490/III_C_R and 13.5% with TAReuk454FWD1/TAReukREV3. Of the remaining reads, 0.1% were singletons or rare sequences (<8 reads) and 46.3% could not be assigned to macrozoobenthic species or to a metazoan phylum. The number of sequences produced in Illumina MiSeq high-throughput sequencing and the number of sequences retained after the main steps of data processing, for each primer-pair in each substrate, are displayed in (Table S.3.1).

From the three artificial substrates sampled along four different sampling times (12 samples), the three primer-pairs were able to identify a total of 171 different taxa, distributed along 9 taxonomic groups: Annelida, Bryozoa, Crustacea, Echinodermata, Hydrozoa, Mollusca, Nemertea, Platyhelminthes and Tunicata (species names and the associated taxonomic classification displayed in Table S.3.2). The highest number of species detected was recorded with the primer-pair TAReuk454FWD1/TAReukREV3 (18S region

– total of 90 species detected), while among the COI-5P primer-pairs, mlCOIintF/LoboR1 retrieved more species than LCO1490/III_C_R (84 species and 63 species, respectively).

The primers employed also differ in their efficiency to recover particular taxonomic groups (Fig. 3.3). However, a non-parametric multivariate analysis indicated no significant statistical differences among the primer-pairs (cophenetic correlation coefficient of the Bray-Curtis similarity = 0.001).

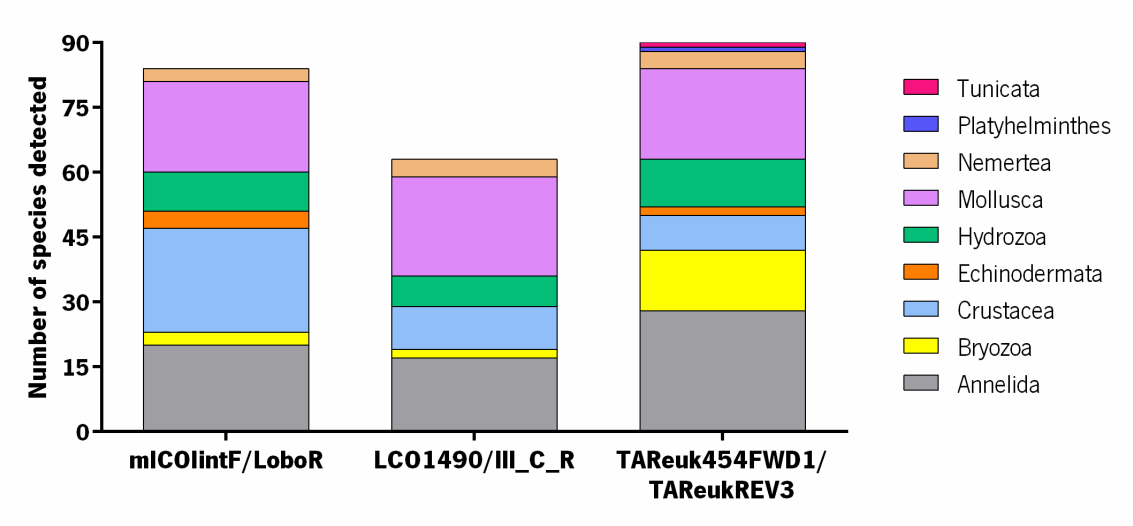


Figure 3.3. Taxonomic profile of the marine macroinvertebrate species detected in the substrates by each primer-pair.

For COI-5P primer-pairs, Crustacea, Mollusca and Annelida were the taxa with higher species diversity (77.3% for mlCOIintF/LoboR1 and 79.7% for LCO1490/III_C_R), while for TAReuk454FWD1/TAReukREV3 the most representative taxonomic groups were Annelida, Mollusca and Bryozoa (70.3% of the total detected taxa; Fig. 3.4). Furthermore, while Tunicata and Platyhelminthes were only detected by 18S, Mollusca and Crustacea had more species identified with COI-primers.

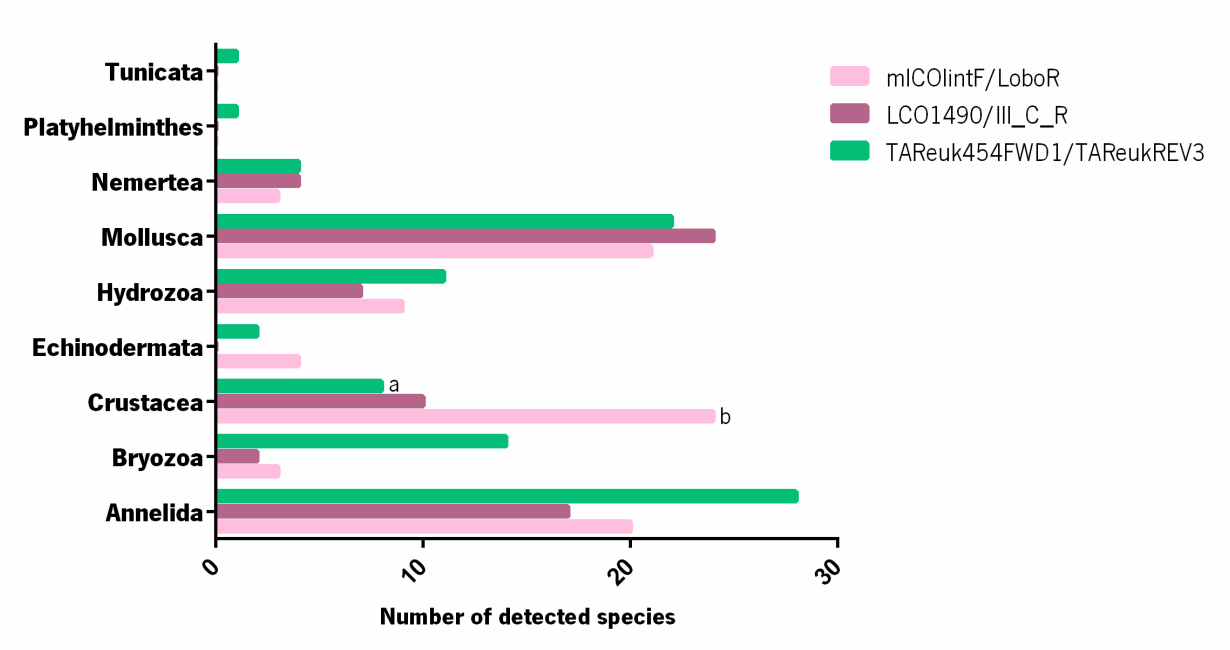


Figure 3.4. Number of species detected, partitioned among the 9 phyla recorded and by primer-pair employed. a and b indicate significant differences between the primers ($p<0.05$).

Regarding to the influence of substrate type and/or sampling time, differences were detected between substrates and primer-pair (cophenetic correlation on coefficient of the Bray-Curtis similarity = 0.55). Consistently, at 7 months was detected the higher species richness for all primer-pairs in the three substrates. Furthermore, as globally before observed, LCO1490/III_C_R also has the lower number of marine macroinvertebrate species detected in all sampling times and substrate types (Fig. 3.5).

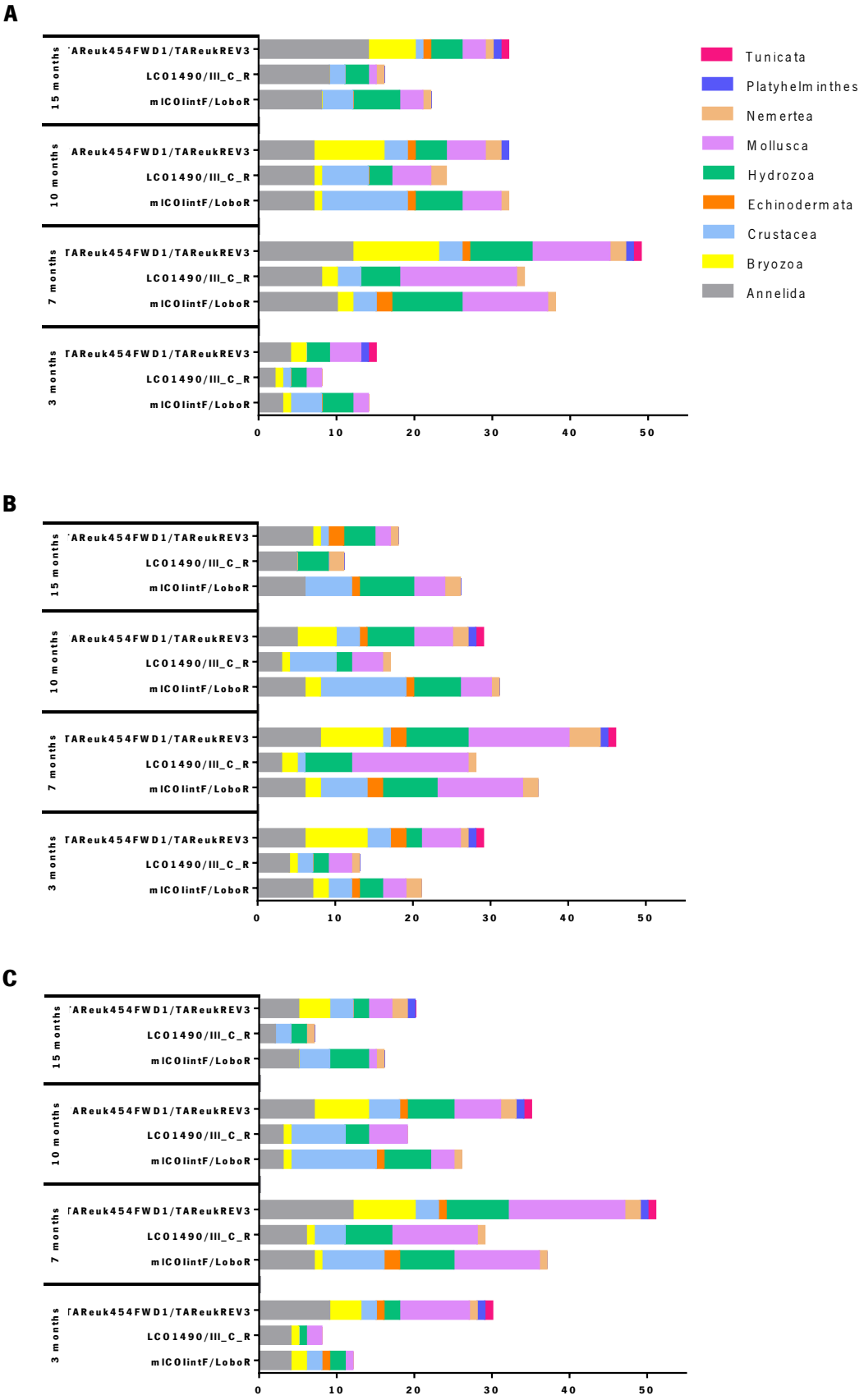


Figure 3.5. Taxonomic composition of marine macroinvertebrate communities for each primer-pair in each substrate type (A - Slate, B - PVC and C - Granite) and sampling time (3, 7, 10 and 15 months).

Combining the detected species by the two COI-5P primer-pairs, 18S V4 region retrieved less taxa than COI (Fig. 3.6.A). However, the three primer-pairs used were highly complementary in their ability to detect marine macroinvertebrate species (Fig. 3.6.B). Among the detected species, only 8.8% were common to the three primer-pairs and 70.1% were exclusively recovered by one primer: 21% for mlCOIinF/LoboR1, 11.7% for LCO1490/Ill_C_R and 37.4% for TAReuk454FWD1/TAReukREV3. For COI-primers, Crustacea (48.6% of the species) was the taxon with more species exclusively detected by mlCOIintF/LoboR1, while for LCO1490/Ill_C_R was Mollusca (55% of the species). For the 18S primer-pairs, 73.4% of the exclusive detected species was Annelida (29.7%), Mollusca (23.4%) and Bryozoa (20.3%).

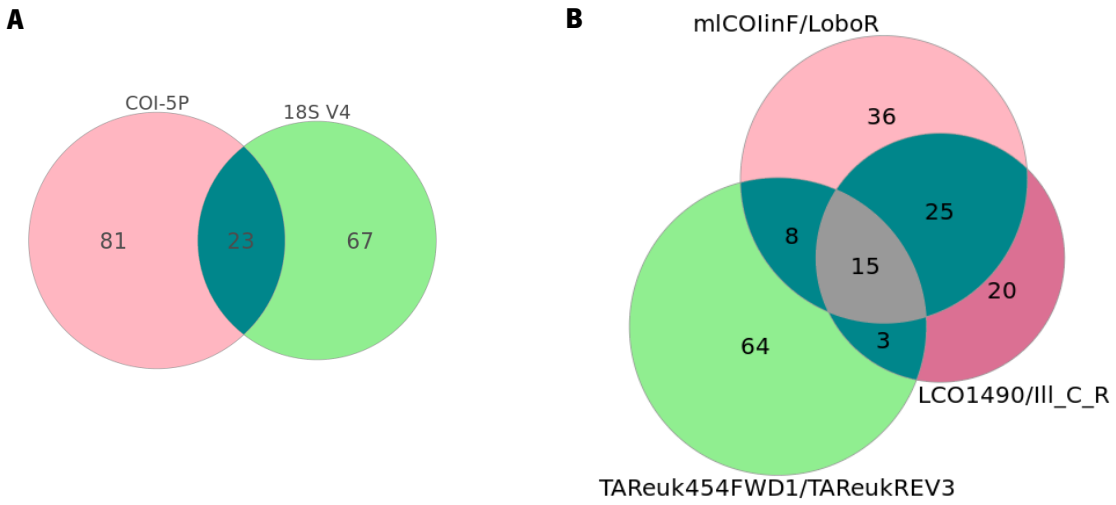


Figure 3.6. Partitioning of the marine macroinvertebrate species detection for (A) both marker loci and (B) primer-pair, in the three substrate types and among all sampling times.

3.3.3 Availability of reference sequences on public databases

From the total 171 marine macroinvertebrate species detected combining the 18S results and the two COI-5P primers considered together, we evaluated the taxonomic coverage in both used databases (mBrave for COI and SILVA for 18S). Gap-analysis showed that 18.7% of the species still lack DNA barcode for COI-5P and 25.7% for 18S rRNA region (Fig. 3.7). While Crustacea was the taxonomic group with higher number of missing sequences in SILVA, in BOLD Bryozoa was the group with more species missing DNA barcodes. No significant statistical differences were detected between the two molecular markers ($p>0.05$).

Based on the comparisons performed among 81 genera detected (Table S.3.3), in 15 there was a higher probability/certainty of the assigned species-level identification, since the genetic distances between all the congeneric species was higher than 3%. In 26 genera, the genetic distance between congeneric species was lower, between 0.5% and 3%. However, after close inspection of matches on SILVA no erroneous taxonomic assignment was detected. In the other genera analyzed, the genetic distances were between 0.2 and 0.5% (5 genus), and only in 8 genera, the genetic distance was zero.

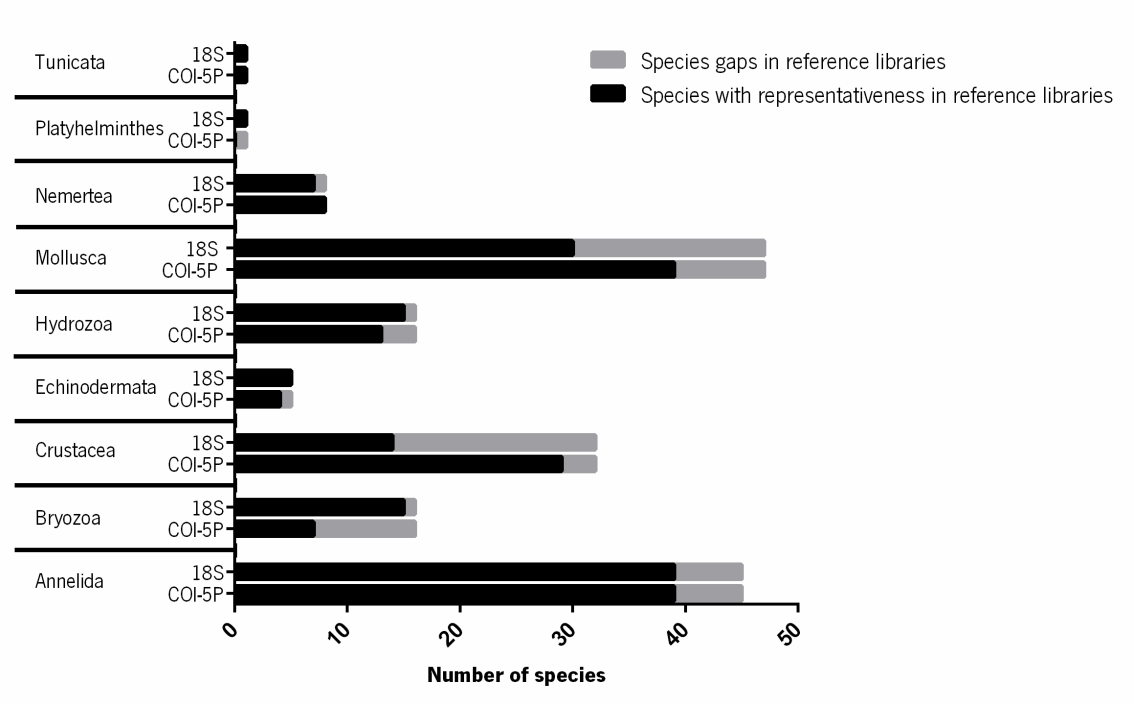


Figure 3.7. Availability of reference sequences of COI-5P and 18S V4 for each taxonomic group of marine macroinvertebrate species detected with the three primer-pairs from COI-5P and 18S genes. Barcode coverage with at least one sequence per species (black bar).

3.4 Discussion

Biomonitoring studies based on DNA metabarcoding approaches require standardized methodologies, namely choice of marker loci and suitable primer-pairs, which can influence the capacity to detect at species-level and characterize the biodiversity present in a marine assemblage. This study revealed the importance of methodological optimization to improve the performance of marine macroinvertebrate species assignments, namely their taxonomic coverage and descrimination, by means of highlighting the influence of amplification process on species detection ability, comparing the taxonomic profiles derived from two molecular markers (COI and 18S rRNA), and revealing marker-specific gaps in species records available in public databases.

In the first screening, we tested the performance of four different COI-5P primer-pairs targeting different segments of the barcode region, where the primers from the 3' region (67.6% and 54.1%) showed higher efficiency compared to primers from 5' region (49.5% and 50.5%). However, the combined results of the top performing primers of the 3' region would produce little gain in terms of species detection success, whereas using the primer-pair mlCOIintF/LoboR1 (3' region) and LCO1490/III_C_R (5' region) we recovered the highest diversity of taxa (respectively 75 + 21 = 96 macroinvertebrate species), corresponding to more than 85% of the total detected species with COI-5P. Therefore, although we observed low to moderate differences in the total number of species detected by each primer-pair, the complementarity between primers is the main criteria to optimize detection success. Primer complementarity possibly resulted from taxon-specific primer-affinity, as reported in previous studies (Hollatz et al., 2017; Zhang et al., 2018). These results emphasize that using a single primer-pair for the COI-5P region will result in fair amount of undetected diversity of marine benthic taxa, therefore we propose the simultaneous employment of the two most complimentary primer-pairs, in this case LCO1490/III_C_R and mlCOIintF/LoboR1 to warrant the highest efficiency in capturing the taxonomic diversity of marine zoobenthic communities.

Comparing mlCOIintF/LoboR1 and mlCOIintF/jgHCO2198, two primer-pairs amplifying exactly the same region of the COI-5P (313 bp at the 3' end) and differing only in the reverse primer, the number of species retrieved was higher with the former (75 and 60, respectively) and the species overlap approximately 67%. A few other studies in which the performance of these two primer-pairs was compared, also indicated slightly better or comparable efficiency of the mlCOIintF/LoboR1 (Haenel et al., 2017; Chang et al., 2020a, b). However, contrary to these results, in a study comparing the performance of the same two primer-pairs,

Derycke et al. (2020) apparently detected a much higher number of species using the primer-pair mlCOIintF/jgHCO2198, leading these authors to recommend only the latter primer-pair for marine invertebrate metabarcoding. Different methodologies adopted in our and Derycke et al. (2020) study may explain the discrepancies in the results. They include the PCR thermal cycling regimes, which had different annealing temperatures, different sequencing depths or the employment of distinct informatics platforms to process HTS reads. Furthermore, taxonomic assignments were performed using different databases (a custom reference library based on BOLD and MIDORI versus unlimited BOLD in our case), which could have different species representativeness and consequently influence species recovery success. It should be also noted that Derycke et al. (2020) does not provide any analyses of the primers' complementarity, in our view a key metric to be considered in primer choice and performance analyses. Another factor to consider in primer choice is the use of several inosine bases in the primer-pair mlCOIintF/jgHCO2198, which greatly increases their cost (Chang et al., 2020a, b) and may affect, or even impede, PCR amplification due to incompatibility with high-fidelity DNA polymerases typically used in the generation of PCR amplicons for HTS metabarcoding (Jungbluth et al., 2020).

For the study of the effect of marker loci on species detection ability, we took into consideration the results of the first screening and selected the two best-performing COI primers (mlCOIintF/LoboR1 and LCO1490/III_C_R), together combined with a primer-pair targeting the V4 region from 18S rRNA gene. The three primer-pairs used were able to detect marine macroinvertebrate taxa, with higher ability to identify at species-level. The amplification of 18S V4 region resulted in more taxa detected than each COI-5P primer-pair individually. However, when compared with the merged results of the two COI primers, 18S retrieved less 8.2% of the species. These results highlight that a best performance of a given marker is not obvious, and in literature there are different and contradictory results, including ours. For example, previous DNA metabarcoding studies using mock zooplankton communities to test a multi-marker (COI and 18S) strategy, demonstrated different taxonomic recovery ability: while one of the studies detected similar patterns of species detection ability among markers (Clarke et al., 2017), the other reported higher detection efficiency with 18S V4 region than with COI, and with higher levels of overlapping between markers (Zhang et al., 2018). However, a direct comparison with our results is limited given the employment of different primers, reference libraries and thresholds. The differences obtained for 18S V4 region in assignment for some taxonomic groups could be related to lower intraspecific variation (Tang et al., 2012; Brown et al., 2015). In addition, taxonomic descrimination of the targeted marker loci, possible mismatches between primer and DNA template caused by PCR bias, or difficulties in primer-affinity for

specific taxa can be possible explanations for the obtained differences between primers and marker region in amplification efficiency.

Our results would suggest one target loci (COI) as the most efficient metabarcoding marker for species assignment on a marine macroinvertebrate community. However, we detected a significant complementarity between the two molecular markers, with each single marker capturing at the very best approximately 61% of the species diversity of a marine community. For example, while isopods were only detected by a COI primer (mlCOIintF/LoboR1), bryozoans were exclusively detected by TAReuk454FWD1/TAReukREV3 (18S). Although few studies compared the performance of molecular marker on species recovery (Dowle et al., 2015; Drummond et al., 2015), and most of them are related to other groups of organisms, our results are generally consistent with previous findings in marine invertebrate communities (Wangensteen et al., 2018), where higher levels of complementarity at species-level between these markers were reported. Since our goal is to capture as widest possible diversity of invertebrate species present in a marine community, and compared to using a single molecular marker, the combination of a multi-locus strategy improves the number of retrieved species, which we recommend as the best practice to be used in marine macroinvertebrates assessments.

We were able to identify most of the macroinvertebrate taxa down to species, even with 18S rRNA gene. We verified the accuracy of species-level assignments, and concluded that the correct species was unambiguously assigned, since the majority of sequences of the detected genus had higher levels of genetic distances. However, some exceptions were detected that probably originated from misidentifications rather than insufficient discrimination ability. For example, the mussel *Mytilus* sp. resulted in multiple sequences attributed to different species identification. The taxonomic assignment uncertainty is typically associated with these bivalve species due to issues related with byparental inheritance of mtDNA, hybridization, divergence between male and female mtDNA genomes, insufficient discrimination ability in some markers and also frequent morphology-based misidentifications (Śmietanka et al., 2004; Araneda et al., 2020; Giusti et al., 2020). Like mentioned before, in such cases the use of a multi-locus strategy could help to detect these cases and solve the taxonomic uncertainties. Moreover, we did not detect any case where two congeneric species had 100% similarity, mostly due to some punctual and probable wrongly identified sequences which had no genetic distance for the species detected by SILVA. These results raise the necessity of quality control and quality assurance tools of the deposited sequences (Fontes et al., 2020).

We performed a comparative gap-analysis of COI and 18S by comparing the species detected in the present study with publicly available sequence records in the respective databases used for taxonomic assignments (respectively BOLD and SILVA). This enabled us to verify if the detection of a species with only one marker could be attributed to gaps in the library of the marker, or, if no gap was found it could be assigned to faulty amplification. A relevant proportion of gaps was recorded for both markers (18.7% for COI and 25.7% for 18S). For example, BOLD do not have representatives of the flatworm *Vorticeros auriculatum*, a species detected by 18S. This incompleteness and inaccuracy of databases may explain some of the species exclusive detected by one marker, and will impede detailed taxonomic discrimination of marine macroinvertebrates, since it closely depends on the taxonomic coverage available on reference libraries (Leray and Knowlton, 2016; Wangensteen et al., 2018). On other hand, some of the detected species with reference sequences in both databases were only detected by one marker (e.g. the tunicate *Asterocarpa humilis*, undetected with COI despite having representative sequences in BOLD). Within the scope of the complete 18S rRNA gene, the selected target region should not be the main reason for failed detections, since V4 is reported to have high amplification success (Brown et al., 2015; Lejzerowicz et al., 2015; Zhang et al., 2018) and demonstrated to have a better performance on taxonomic assignments than other 18S regions (Fais et al., 2020b). However, since primers demonstrated to fail in some species detected, the obtained results were probably due to faulty amplification, which emphasize the necessity of the use of at least two markers and multiple primers. Although the availability of sequences was not the main influence for species detection, it revealed that more investment should be allocated to obtain reliable reference sequences to enhance species assignment accuracy, in order to achieve taxonomic profile of a target community as complete as possible. Indeed, although COI-based monitoring approaches may claim the advantage to having a verified and dedicated database for a large variety of taxa, several studies already reported the existence of significant gaps in reference libraries particularly for marine macrobenthic taxa (Weigand et al., 2019: 30% to 50% of completion of databases), including for the region that comprised the study area targeted in the current study (Leite et al., 2020: 37% of taxonomic coverage).

The three primer-pairs used in this study were able to detect marine macroinvertebrate species in every sampling time, all of them consistently pointing to a higher species diversity after 7 months of deployment of the substrates. These results highlight the benefit of the application of a multiple primer-pair and multi-locus strategy for ecological assessments of marine species, since if we had only used one primer-pair or marker we would have failed to detect important macrobenthic taxa, and the taxonomic profile of the community could emerge substantially different. Temporal and seasonal changes in a community could

affect the potential of species monitoring, especially when methodological bias originated by amplification procedures (choice of marker loci and primer-pairs) could influence ecological interpretations (Clarke et al., 2017).

These results were an evidence of the influence of marker choice on the ability of macrobenthic species identification and the necessity to overcome the drawback of reference libraries incompleteness. Even the proportion of gaps was not a dominant problem for the detected species, with more completed reference libraries we probably could reduce the number of reads without taxonomic assignment. For future high-throughput assessments using DNA metabarcoding approaches, we recommended combining molecular markers, and if possible multiple primer-pairs, to increase the accuracy of species-level detection and biodiversity estimation, to overcome taxonomic gaps resulting from primer bias, and to yield reliable results for marine macroinvertebrate monitoring studies.

References

Alberdi A., Aizpurua O., Gilbert M.T.P., Bohmann K. (2017). Scrutinizing key steps for reliable metabarcoding of environmental samples. Methods in Ecology and Evolution. 9(1), 134-147. https://doi.org/10.1111/2041-210X.12849.

Andújar C., Arribas P., Yu D.W., Vogler A.P., Emerson B.C. (2018). Why the COI barcode should be the community DNA metabarcode for the metazoa. Molecular Ecology. 27, 3968 –3975. https://doi.org/10.1111/mec.14844.

Araneda T.C., Pardo M.Á., Jimenez E., Longa Á., Lee R., Segura C., Zbawicka M., Wenne R., Gardner J.P., Larraín M.A. (2020). A comment on comment on Giusti et al. (2020) "Mussels (*Mytilus* spp.) products authentication: A case study on the Italian market confirms issues in species identification and arises concern on commercial names attribution. Food Control. 121, 107626. https://doi.org/10.1016/j.foodcont.2020.107626.

Aylagas E., Borja Á., Muxika I., Rodríguez-Ezpeleta N. (2018). Adapting metabarcoding-based benthic biomonitoring into routine marine ecological status assessment networks. Ecological Indicators. 95, 194-202. https://doi.org/10.1016/j.ecolind.2018.07.044.

Bista I., Carvalho G.R., Tang M., Walsh K., Zhou X., Hajibabaei M., Shokralla S., Seymour M., Bradley D., Liu S., Christmas M., Creer S. (2018). Performance of amplicon and shotgun sequencing for accurate biomass estimation in invertebrate community samples. Molecular Ecology Resources. 18, 1020–1034. https://doi.org/10.1111/1755-0998.12888.

Braukmann T.W.A., Ivanova N.V., Prosser S.W.J., Elbrecht V., Steinke D., Ratnasingham S., deWaard, J.R., Sone J.E., Zakharov E.V., Hebert P.D.N. (2019). Metabarcoding a diverse arthropod mock community. Molecular Ecology Resources. 19, 711–727. https://doi. org/10.1111/1755-0998.13008.

Brown E.A., Chain F.J.J., Crease T.J., MacIsaac H.J., Cristescu M.E. (2015). Divergence thresholds and divergent biodiversity estimates: Can metabarcoding reliably describe zooplankton communities? Ecology and Evolution. 5, 2234–2251. https://doi.org/10.1002/ ece3.1485.

Chamberlain S. (2019). bold: Interface to bold systems API. R package version 0.9.0. https://CRAN.R-project.org/package=bol.

Chang J.J.M., Ip Y.C.A., Bauman A.G., Huang D. (2020a). MinION-in-ARMS: Nanopore sequencing to expedite barcoding of specimen-rich macrofaunal samples from Autonomous Reef Monitoring Structures. Frontiers in Marine Science. 7, 448. https://doi.org/10.3389/fmars.2020.00448.

Chang J.J.M., Ip Y.C.A., Ng C.S.L., Huang D. (2020b). Takeaways from mobile DNA barcoding with bentoLab and MinION. Genes. 11(10), 1121. https://doi.org/10.3390/genes11101121.

Clarke K.R., Gorley R.N. (2006). PRIMER V6: user manual/tutorial. PRIMER-E Ltd, Plymouth.

Clarke L.J., Beard J.M., Swadling K.M., Deagle B.E. (2017). Effect of marker choice and thermal cycling protocol on zooplankton DNA metabarcoding studies. Ecology and Evolution. 7(3), 873-883. https://doi.org/10.1002/ece3.2667.

Clarke L.J., Soubrier J., Weyrich L.S., Cooper A. (2014). Environmental metabarcodes for insects: *in silico* PCR reveals potential for taxonomic bias. Molecular Ecology Resources. 14, 1160–1170. https://doi.org/10.1111/1755-0998.12265.

Collins R.A., Bakker J., Wangensteen O.S., Soto, A.Z., Corrigan L., Sims D.W., Genner M.J., Mariani S. (2019). Non-specific amplification compromises environmental DNA metabarcoding with COI. Methods in Ecology and Evolution. 10(11), 1985-2001. https ://doi.org/10.1111/2041-210X.13276.

Cowart D.A., Pinheiro M., Mouchel O., Maguer M., Grall J., Miné J., Arnaud-Haond S. (2015). Metabarcoding is powerful yet still blind: A comparative analysis of morphological and molecular surveys of seagrass communities. PLoS one. 10(2), e0117562. https ://doi.org/10.1371/journal. pone.0117562.

Deagle B.E., Jarman S.N., Coissac E., Pompanon F., Taberlet P. (2014). DNA metabarcoding and the cytochrome c oxidase subunit I marker: Not a perfect match. Biology Letters. 10(9), 20140562. https://doi.org/10.1098/rsbl.2014.0562.

Derycke S., Maes S., van den Bulcke L., Vanhollebeke J., Wittoeck J., Hillewaert H., Ampe B., Haegeman A., Hostens K., de Backer A. (2020). Optimisation of metabarcoding for monitoring marine macrobenthos: primer choice and morphological traits determine species detection in bulkDNA and eDNA from the ethanol preservative. https://doi.org/10.22541/au.159665093.39993653.

Dowle E.J., Pochon X., Banks J., Shearer K., Wood S.A. (2015). Targeted gene enrichment and high throughput sequencing for environmental biomonitoring: a case study using freshwater macroinvertebrates. Molecular Ecology Resources. 16(5), 1240–1254. https://doi.org/10.1111/1755-0998.12488.

Drummond A.J., Newcomb R.D., Buckley T.R., Xie D., Dopheide A., Potter B.C.M., Heled J., Ross H.A., Tooman L., Grosser S., Park D., Demetras N.J., Stevens M.I., Russell J.C., Anderson S.H., Carter A., Nelson N. (2015). Evaluating a multigene environmental DNA approach for biodiversity assessment. Gigascience. 4(1), s13742-015. https://doi.org/10.1186/s13742-015-0086-1.

Duarte S., Vieira P.E., Lavrador A.S., Costa F.O. (2021). Status and prospects of marine NIS detection and monitoring through (e)DNA metabarcoding. Science of the Total Environmental. 751, 141729. https://doi.org/10.1016/j.scitotenv.2020.141729.

Elbrecht V., Braukmann T.W.A., Ivanova N.V., Prosser S.W.J., Hajibabaei M., Wright M., Zakharov E.V., Hebert P.D.N., Steinke D. (2019). Validation of COI metabarcoding primers for terrestrial arthropods. PeerJ. 7, e7745. http://doi.org/10.7717/peerj.7745.

Elbrecht V., Leese F. (2017). Validation and development of COI metabarcoding primers for freshwater macroinvertebrate bioassessment. Frontiers in Environmental Science. 5, 1–11. https://doi.org/10.3389/fenvs.2017.00011.

Fais M., Bellisario B., Duarte S., Vieira P.E., Sousa R., Canchaya C., Costa F.O. (2020a). Meiofauna metabarcoding in Lima estuary (Portugal) suggests high taxon replacement within a background of network stability. Regional Studies in Marine Science. 38, 101341. https://doi.org/10.1016/j.rsma.2020.101341.

Fais M., Duarte S., Vieira P.E., Sousa R., Hajibabaei M., Canchaya C., Costa F.O. (2020b). Small-scale spatial variation of meiofaunal communities in Lima estuary (NW Portugal) assessed through metabarcoding. Estuarine, Coastal and Shelf Science. 106683. https://doi.org/10.1016/j.ecss.2020.106683.

Flynn J.M., Brown E.A., Chain F.J.J., MacIsaac H.J., Cristescu M.E. (2015). Toward accurate molecular identification of species in complex environmental samples: testing the performance of sequence filtering and clustering methods. Ecology and Evolution. 5, 2252–2266. https://doi.org/10.1002/ece3.1497.

Folmer O., Black M., Hoeh W., Lutz R., Vrijenhoek R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology. 3(5), 294–9.

Fontes J.T., Vieira P.E., Ekrem T., Soares P., Costa F.O. (2020). BAGS: an automated Barcode, Audit & Grade System for DNA barcode reference libraries. Molecular Ecology Resources. https://doi.org/10.1111/1755-0998.13262.

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, 851–861. https://doi.org/10.1111/1755-0998.12138.

Gibson J.F., Shokralla S., Curry C., Baird D.J., Monk W.A., King I., Hajibabaei M. (2015). Large-scale biomonitoring of remote and threatened ecosystems via high-throughput sequencing. PLoS One. 10(10), e0138432. https://doi.org/10.1371/journal.pone.0138432.

Giebner H., Langen K., Bourlat S.J., Kukowka S., Mayer C., Astrin J.J. Misof B., Fonseca V.G. (2020). Comparing diversity levels in environmental samples: DNA sequence capture and metabarcoding approaches using 18S and COI genes. Molecular Ecology Resources. 20(5), 1333-1345. https://doi.org/10.1111/1755-0998.13201.

Giusti A., Tosi F., Tinacci L., Guardone L., Corti I., Arcangeli G., Armani A. (2020). Mussels (*Mytilus* spp.) products authentication: A case study on the Italian market confirms issues in species identification and arises concern on commercial names attribution. Food Control. 107379. https://doi.org/10.1016/j.foodcont.2020.107379.

Haenel Q., Holovachov O., Jondelius U., Sundberg P., Bourlat S.J. (2017). NGS-based biodiversity and community structure analysis of meiofaunal eukaryotes in shell sand from Hållö island, Smögen, and soft mud from Gullmarn Fjord, Sweden. Biodiversity Data Journal. https://doi.org/10.3897/BDJ.5.e12731.

Hajibabaei M., Porter T.M., Wright M., Rudar J. (2019). COI metabarcoding primer choice affects richness and recovery of indicator taxa in freshwater systems. PLoS One. 14, e0220953. https://doi.org/10.1371/journal.pone.0220953.

Hebert P.D.N., Cywinska A., Ball S.L., deWaard J.R. (2003). Biological identifications through DNA barcodes. Proceedings of the Royal Society of London. Series B: Biological Sciences. 270(1512), 313-321. https://doi.org/10.1098/rspb.2002.2218.

Hollatz C., Leite B.R., Lobo J., Froufe H., Egas C., Costa F.O. (2017). Priming of a DNA metabarcoding approach for species identification and inventory in marine macrobenthic communities. Genome. 60(3), 260-271. https://doi.org/10.1139/gen-2015-0220.

Illumina. (2013). 16S Metagenomic sequencing library preparation preparing 16S Ribosomal RNA gene amplicons for the Illumina MiSeq system.

Ip Y.C.A., Tay Y.C., Gan S.X., Ang H.P., Tun K., Chou L.M., Huang D., Meier R. (2019). From marine park to future genomic observatory? Enhancing marine biodiversity assessments using a biocode approach. Biodiversity Data Journal. 7, e46833. https://doi.org/10.3897/BDJ.7.e46833.

Ivanova N.V., deWaard J.R., Hebert P.D.N. (2006). An inexpensive, automation-friendly protocol for recovering high-quality DNA. Molecular Ecology Notes. 6, 998–1002. https://doi.org/10.1111/j.1471-8286.2006.01428.x.

Jungbluth M.J., Burns J., Grimaldo L., Slaughter A., Katla A., Kimmerer W. (2020). Feeding habits and novel prey of larval fishes in the northern San Francisco Estuary. BioRxiv. https://doi.org/10.1101/2020.10.18.344440.

Kebschull J.M., Zador A.M. (2015). Sources of PCR-induced distortions in high-throughput sequencing data sets. Nucleic Acids Research. 43(21), e143. https://doi.org/10.1093/nar/gkv717.

Kozich J.J., Westcott S.L., Baxter N.T., Highlander S.K., Schloss P.D. (2013). Development of a dualindex sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumine sequencing platform. Applied and Environmental Microbiology. 79, 5112–5120. https://doi.org/10.1128/AEM.01043-13.

Kumar S., Stecher G., Tamura K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Molecular Biology and Evolution. 33(7), 1870-1874. https://doi.org/10.1093/molbev/msw054.

Lacoursière-Roussel A., Howland K., Normandeau E., Grey E.K., Archambault P., Deiner K., Lodge D.M., Hernandez C., Leduc N., Bernatchez L. (2019). eDNA metabarcoding as a new surveillance approach for coastal Arctic biodiversity. Ecology and Evolution. 8, 7763–7777. https://doi.org/10.1002/ece3.4213.

Leite B.R., Vieira P.E., Teixeira M.A.L., Lobo-Arteaga J., Hollatz C., Borges L.M.S., Duarte S., Troncoso J.S., Costa F.O. (2020). Gap-analysis and annotated reference library for supporting macroinvertebrate metabarcoding in Atlantic Iberia. Regional Studies in Marine Science. 36, 101307. https://doi.org/10.1016/j.rsma.2020.101307.

Lejzerowicz F., Esling P., Pillet L., Wilding T.A., Black K.D., Pawlowski J. (2015). High-throughput sequencing and morphology perform equally well for benthic monitoring of marine ecosystems. Scientific Reports. 5, 13932. https://doi.org/10.1038/srep13932.

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. 112(7), 2076–2081. https://doi.org/10.1073/pnas.1424997112.

Leray M., Knowlton N. (2016). Censusing marine eukaryotic diversity in the twenty-first century. Philosophical Transactions of the Royal Society. B. 371, 20150331. http://dx.doi.org/10.1098/rstb.2015.0331.

Leray M., Knowlton N. (2017). Random sampling causes the low reproducibility of rare eukaryotic OTUs in Illumina COI metabarcoding. PeerJ. 5, e3006. https://doi.org/10.7717/peerj.3006.

Leray M., Yang J.Y., Meyer C.P., Mills S.C., Agudelo N., Ranwez V., Boehm J.T., 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(1), 1–14. https://doi.org/10.1186/1742-9994-10-34.

Lobo J., Costa P.M., Teixeira M.A., Ferreira M.S., Costa M.H.C., Costa F.O. (2013). Enhanced primers for amplification of DNA barcodes from a broad range of marine metazoans. BMC Ecology. 13(1), 34. https://doi.org/10.1186/1472-6785-13-34.

Lobo J., Shokralla S., Costa M.H., Hajibabaei M., Costa F.O. (2017). DNA metabarcoding for high-throughput monitoring of estuarine macrobenthic communities. Scientific Reports. 7, 15618. https://doi.org/10.1038/s41598-017-15823-6.

McGee K.M., Robinson C.V., Hajibabaei M. (2019). Gaps in DNA-based biomonitoring across the globe. Frontiers in Ecology and Evolution. 7, 337. https://doi.org/10.3389/fevo.2019.00337.

Piñol J., Mir G., Gomez-Polo P., Agusti N. (2015). Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. Molecular Ecology Resources. 15, 819–830. https://doi.org/10.1111/1755-0998.12355.

Piñol J., Senar M.A., Symondson W.O.C. (2019). The choice of universal primers and the characteristics of the species mixture determine when DNA metabarcoding can be quantitative. Molecular Ecology. 28, 407–419. https://doi.org/10.1111/mec.14776.

Porter T.M., Hajibabaei M. (2020). Putting COI metabarcoding in context: the utility of Exact Sequence Variants (ESVs) in biodiversity analysis. Frontiers in Ecology and Evolution. 8, 248. https://doi.org/10.3389/fevo.2020.00248.

Prego R., Fraga F. (1992). A simple model to calculate the residual flows in a Spanish ria. Hydrographic consequences in the ria of Vigo. Estuarine Coastal and Shelf Science. 34, 603e615. http://doi.org/10.1016/S0272-7714(05)80065-4.

Pruesse E., Quast C., Knittel K., Fuchs B.M., Ludwig W., Peplies J., Glöckner F.O. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Research. 35(21), 7188-7196. https://doi.org/10.1093/nar/gkm864.

Quast C., Pruesse E., Yilmaz P., Gerken J., Schweer T., Yarza P., Peplies J., Glockner F.O. (2013). The SILVA ribossomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research. 41(D1), D590-D596: https://doi.org/10.1093/nar/gks1219.

Ransome E., Geller J.B., Timmers M., Leray M., Mahardini A., Sembiring A., Collins A.G., 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, e0175066. https://doi.org/10.1371/journal.pone.0175066.

Ratnasingham S. (2019). mBRAVE: The Multiplex Barcode Research and Visualization Environment.Biodiversity Information Science and Standards. 3, e37986. https://doi.org/10.3897/biss.3.37986.

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

Schloss P.D., Westcott S.L., Ryabin T., Hall J.R., Hartmann M., Hollister E.B., Lesniewski R.A., Oakley B.B., Parks D.H., Robinson C.J., Sahl J.W., Stres B., Thallinger G.G., van Horn D.J., Weber C.F. (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and Enviromental. Microbiology. 75(23), 7537-7541. https://doi.org/10.1128/AEM.01541-09.

Schmieder R., Edwards R. (2011). Quality control and preprocessing of metagenomic datasets. Bioinformatics. 27(6), 863–4. https://doi.org/10.1093/bioinformatics/btr026.

Schubert M., Lindgreen S., Orlando L. (2016). AdapterRemoval v2: rapid adapter trimming, identification, and read merging Findings Background. BMC Research Notes. 9(1), 1-7. https://doi.org/10.1186/s13104-016-1900-2.

Shokralla S., Porter T.M., Gibson J. F., Dobosz R., Janzen D.H., Hallwachs,W., Golding G.B., Hajibabaei M. (2015). Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform. Scientific Reports. 5, 9687. https://doi.org/10.1038/srep09687.

Śmietanka B., Zbawicka M., Wolowicz M., Wenne R. (2004). Mitochondrial DNA lineages in the European populations of mussels (*Mytilus* spp.). Marine Biology. 146, 79–92. https://doi.org/10.1007/s00227-004-1418-3.

Stoeck T., Bass D., Nebel M., Christen R., Jones M.D.M., Breiner H.W., Richards T.A. (2010). Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. Molecular Ecology. 19, 21–31. https://doi.org/10.1111/j.1365-294X.2009.04480.x.

Taberlet P., Coissac E., Pompanon F., Brochmann C., Willerslev E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. Molecular Ecology. 21(8), 2045–2050. https://doi.org/ 10.1111/j.1365-94X.2012.05470.x.

Tang C.Q., Leasi F., Obertegger U., Kieneke A., Barraclough T.G., 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. 109(40), 16208–16212. https://doi.org/10.1073/ pnas.12091 60109.

Thompson J.D., Higgins D.G., Gibson T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research. 22(22), 4673-4680. https://doi.org/10.1093/nar/22.22.4673.

van der Loos L.M., Nijland R. (2020). Biases in bulk: DNA metabarcoding of marine communities and the methodology involved. Molecular Ecology. https://doi.org/10.1111/mec.15592.

Veiga P., Torres A.C., Aneiros F., Sousa-Pinto I., Troncoso J.S., Rubal M. (2016). Consistent patterns of variation in macrobenthic assemblages and environmental variables over multiple spatial scales using taxonomic and functional approaches. Marine Environmental Research. 120, 191-201. https://doi.org/10.1016/j.marenvres.2016.08.011.

Vieira P.E., Desiderato A., Holdich D.M., Soares P., Creer S., Carvalho G.R., Costa F.O., Queiroga H. (2019). Deep segregation in the open ocean: Macaronesia as an evolutionary hotspot for low dispersal marine invertebrates. Molecular Ecology. 28 (7), 1784–1800. http://doi.org/10.1111/mec.15052.

von Ammon U., Wood S.A., Laroche O., Zaiko A., Tait L., Lavery S., Inglis G.J., Pochon X. (2018). Combining morpho-taxonomy and metabarcoding detection of non-indigenous marine pests in biofouling communities. Scientific Reports. 8(1), 1-11. https://doi.org/10.1038/s41598-018-34541-1.

Wangensteen O.S., Palacín C., Guardiola M., Turon X. (2018). DNA metabarcoding of littoral hard-bottom communities: high diversity and database gaps revealed by two molecular markers. PeerJ. 6, e4705. http://doi.org/10.7717/peerj.4705.

Weigand H., Beermann A.J., Ciampor F., Costa F.O., Csabai Z., Duarte S., Geiger M.F., Grabowski M., Rimet F., Rulik B., Strand M., Szucsich N., Weigand A.M., Willassen E., Wyler S.A., Bouchez A., Borja A., Ciamporová-Zaťovicová Z., Ferreira S., Dijkstra K.B., Eisendle U., Freyhof J., Gadawski P., Graf W., Haegerbaeumer A., van der Hoorn B.B., Japoshvili B., Keresztes L., Keskin E., Lesse F., Macher J.N., Mamos T., Paz G., Pesic V., Pfannkuchen D.M., Pfannkuchen M.A., Price B.W., Rinkevich B., Teixeira M.A.L., Várbíró G., Ekrem T. (2019). DNA barcode reference libraries for the monitoring of aquatic biota in Europe: Gap-analysis and recommendations for future work. Science of the Total Environment. 678, 499–524. http://doi.org/10.1016/j.scitotenv.2019.04.247.

WoRMS Editorial Board (2020). World Register of Marine Species. Available from http://www.marinespecies.org at VLIZ. Accessed 2020-07-28. http://doi.org/10.14284/170.

Zaiko A., Samuiloviene A., Ardura A., Garcia-Vazquez E. (2015). Metabarcoding approach for nonindigenous species surveillance in marine coastal waters. Marine Pollution Bulletin. 100, 53–59. https://doi.org/10.1016/j.marpolbul.2015.09.030.

Zhang G.K., Chain F.J., Abbott C.L., Cristescu M.E. (2018). Metabarcoding using multiplexed markers increases species detection in complex zooplankton communities. Evolutionary Applications. 11(10), 1901-1914. https://doi.org/10.1111/eva.12694.

Chapter 4

Coastal macrozoobenthos monitoring: impact of artificial substrate material and comparison of morphology vs DNA metabarcoding-based taxonomic profiling

Abstract

Large-scale marine biodiversity monitoring remains challenging because it is commonly based on morphological identification, which is time-consuming and requires specialized expertise. DNA metabarcoding has high potential to improve our knowledge on marine macroinvertebrate biodiversity. Artificial substrates deployed in coastal areas may selectively influence and enhance the colonization of different species, and coupled with high-throughput techniques (i.e. DNA metabarcoding) provide the ability to accurately assess, manage and monitor marine communities. Here, we deployed three types of artificial substrates at Toralla Island (NW Iberian Peninsula) for 3, 7, 10 and 15 months to test the influence of different types of artificial substrates on the composition of marine macrozoobenthic communities. We combined morphology and DNA metabarcoding (COI and 18S rRNA genes) to compare the community composition and diversity patterns and to evaluate the ability of both methods to detect and enable monitoring of coastal macrozoobenthic colonization. DNA metabarcoding retrieved more taxa with higher ability to identify at species-level, even though some of the species were detected only using morphology. Experimental design, sample processing, incompleteness of reference databases or DNA-methodologies can explain these findings. The total number of species detected was slightly different among substrates, and after 7 months we detected the highest number of species. However, some taxa revealed different temporal-patterns of colonization, as well as variations in taxonomic diversity among substrates and seasons. In general, using artificial substrates which promote macrozoobenthic colonization, coupled with DNA-based identification approaches capture the composition of a marine macroinvertebrate community within the recognized biogeographical scope for the study area. Slate substrate harbored a higher number of species and sampling times influenced the recruitment of macrozoobenthos. DNA metabarcoding and morphology were somewhat complementary in their ability to detect macrozoobenthic species. Although DNA metabarcoding efficiency is expected to improve, currently we propose that both approaches should be used whenever feasible to avoid missing relevant taxa. The results highlight the potential of the adopted experimental strategy to be used in coastal biomonitoring programs targeting macroinvertebrate biodiversity.

Keywords: DNA metabarcoding; Morphology; Marine macrozoobenthos; Artificial substrates; Ecological succession; Community analysis.

4.1 Introduction

Marine biodiversity has a longer wavelength on ecological and evolutionary time scales, being a genetically privileged ecosystem composed by highly diverse genetic material (Kelly et al., 2014). The establishment and diversity of species present in a marine community is affected by the levels of perturbations and the recruitment ability, which is influenced by timing, availability of colonizers, abundance of species and space (Bowden et al., 2006; Wahl et al., 2011). To understand the processes of settlements in macroinvertebrate communities is essential to elucidate the dynamics in a benthic assemblage and to recognize spatio-temporal patterns of colonization (Sokolowski et al., 2017).

The difficulty in finding standardized approaches to routinely and accurately sample marine macroinvertebrate assemblages makes it hard to compare the diversity of species present in the community and to understand patterns of variation. Artificial substrates deployed in selected locations may selectively enhance the development of benthic communities, providing a solution for improving standardization and replicability (Edgar and Klumpp, 2003). The choice of artificial substrates is commonly based on the cost-effectiveness of materials: local availability, durability (resistant to chemical and physical factors) and non-destructive characteristics (Field et al., 2007; Spagnolo et al., 2014), as well as on the deployment method (orientation, light influence, hydrodynamics, wave exposure) (Pacheco et al., 2010). All these variables have influence on species colonization ability, affecting the settlement and development of macroinvertebrate communities (Moura et al., 2008; Antoniadou et al., 2010). Although ecological studies have been used artificial substrates to assess biodiversity and to investigate their impact on species colonization (Gee and Warwick, 1996; Cangussu et al., 2010; Spagnolo et al., 2014; Cacabelos et al., 2016, 2020; Mallela et al., 2017; Marraffini et al., 2017; Sokolowski et al., 2017), the processes of colonization and succession are not clearly understood (Underwood and Chapman, 2006).

The short or long-term monitoring of a community will be determined by the ability of the species to colonize the substrates in response to temporal changes (i.e. succession) (García-Sanz et al., 2014). Some ecological factors such as the early species present in the ecosystem at the time of deployment, the influence of interactions for space, the indirect effects of consumers will create the community structure and therefore establish a complex macroinvertebrate assemblage.

Marine macroinvertebrate communities are composed by a very diverse set of species belonging to a wide range of phyla, varying in structure over time (Thrush et al., 1994). These communities have been widely used to assess responses to environmental disturbances and the ecological status of marine and estuarine

habitats (Thrush and Dayton, 2002). The identification of macroinvertebrate species requires a reliable and robust method for standard biodiversity monitoring (Cristescu, 2014). Biodiversity assessment is commonly performed through morphology-based species identification, which provides data on species occurrences and abundances. However, this approach is time-consuming, expensive and require taxonomic specialists (Yu et al., 2012), and is particularly difficult in marine communities due to the phylogenetic diversity of the target taxa.

DNA-sequencing technologies have seen improvements through the implementation of high-throughput sequencing (HTS), which, coupled with DNA barcoding (i.e. short DNA sequences as a molecular tag for species identification; Hebert et al., 2003), are making DNA metabarcoding (Taberlet et al., 2012) becoming one of the most efficient, rapid and cost-effective genomic approach to assess species biodiversity for biomonitoring programs of marine macroinvertebrate communities (Dowle et al., 2016; Carew et al., 2018). Contrary to morphology, DNA metabarcoding allow the identification of small fragments of organisms, reveals hidden diversity and distinguish cryptic species, including among marine macroinvertebrates (Lindeque et al., 2013; Pearman et al., 2018; Carvalho et al., 2019), and it also allows studies of large spatial and temporal scales (Leray and Knowlton, 2015).

Several studies have already compared morphology and DNA-based identification approaches to assess marine macroinvertebrate diversity (Kelly et al., 2017; Lobo et al., 2017; Aylagas et al., 2018; von Ammon et al., 2018; Schroeder et al., 2020; Steyaert et al., 2020). However, few studies used different target genetic regions for metabarcoding, and more than one primer combination, and few analyze short- and long-term of temporal variation in these communities to study processes of colonization. Although comparing DNA metabarcoding results to morphological identifications are difficult mostly due to the different levels of taxonomic identification obtained (Cahill et al., 2018), further research must be undertaken to compare morphology-based and DNA metabarcoding identifications, to determine differences and complementarity between both approaches.

In this study, we use three types of artificial substrates (slate, PVC and granite) to evaluate their impact on the composition and structure of marine macrozoobenthic communities. In addition, we evaluate the short- and long-term and seasonal patterns of macrobenthic colonization in NW Atlantic Iberia. We combined DNA metabarcoding (COI and 18S rRNA genes) and morphology to assess the diversity and community composition of marine macrozoobenthos. We also compare the ability of both methods to detect and enable monitoring of coastal macrozoobenthic colonization.

4.2 Materials and methods

4.2.1 Sampling design

In this study we used the same substrates types (slate, PVC and granite) and deployed those at the same location (Toralla Island, NW Iberian Peninsula; Fig. 4.1) as performed in Chapter 3 (please see 3.3 Materials and Methods section for more details).

Sixteen replicates of each of the three selected artificial substrates (Fig. 4.1) were deployed in random order in December 2016 (Fig. S.4.1). After 3 months (March 2017), 7 months (July 2017), 10 months (October 2017) and 15 months (March 2018) four replicates of each substrate were randomly chosen and removed. A hermetic plastic bag was used to collect the replicates, which were individually photographed in the laboratory (see Fig. 4.2). Three replicates of each substrate were used for morphological identification and one for DNA metabarcoding.

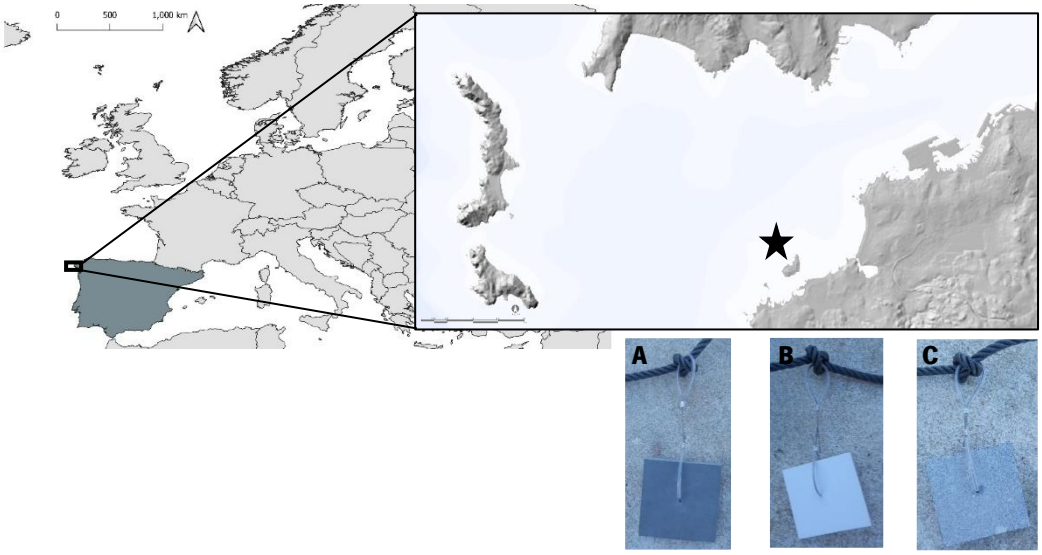


Figure 4.1. Map of the Iberian Peninsula[2] showing the sampling site (black star: Toralla Island, NW Iberian Peninsula) and the three artificial substrates (10 x 10 cm) used for marine macrozoobenthos colonization. A – Slate; B – PVC; C – Granite.

[2] Sousa-Guedes D., Arenas-Castro S., Sillero N. (2020). Ecological niche models reveal climate change effect on biogeographical regions: the Iberian Peninsula as a case study. Climate. 8(3), 42. https://doi.org/10.3390/cli8030042.

4.2.2 Morphology-based taxonomic identification

Each substrate was carefully washed, and using a brush and tweezers mobile fauna was separated from sessile. The sessile fauna was identified during sample processing, and the mobile fauna was sieved using a 500 μm mesh. The samples from 3 and 7 months were preserved in 4% formaldehyde, and the samples from 9 and 15 months were preserved in ethanol, and stored at -20 ^{0}C. Identifications were performed under a binocular microscope and with the help of taxonomic identification manuals and keys (e.g. Naylor, 1972; Lincoln, 1979; Conlan, 1990; Hayward and Ryland, 1995; Saldanha, 1995; Guerra-García et al., 2013; Gouillieux and Sobre, 2015). Morphological identifications of the specimens were carried out to the lowest possible taxonomic level, and the number of individuals belonging to each taxonomic group was counted. The validity of the species names was then verified in the World Register of Marine Species (WoRMS) database.

Three substrate replicates were lost from the 15 months sampling time, one of slate and two of granite, probably due to strong marine currents that broke the cables where they were suspended, and could not be included in further analysis.

4.2.4 DNA metabarcoding-based taxonomic identification

For DNA metabarcoding analysis, each substrate was carefully washed using filtered seawater and the mobile and sessile fauna were processed separately. First, the mobile fauna was brushed and sieved using a 500 μm mesh, and then the sessile fauna was scraped with a spatula into a tray. The water in the washing container was sieved and preserved with mobile fauna. All samples were then preserved in ethanol and stored at -20 ^{0}C until further analysis.

4.2.4.1 DNA extraction, PCR amplification, HTS procedures and data processing

DNA metabarcoding procedures were performed as in Chapter 3 (please see 3.3.2.1 on Materials and Methods section for more details). A lysis buffer solution was used to obtain the lysates from each substrate/sampling time combination, and DNA extraction procedures were adapted from Ivanova et al. (2006), and for each sample we duplicate the lysates. Two primer-pairs targeting the COI-5P barcode region (mlCOIintF/LoboR1 and LCO1490/III_C_R) and one primer-pair targeting the V4 region of the 18S rRNA gene (TAReuk454FWD1/TAReukREV3) were selected for PCR amplification of DNA templates of the

marine macrozoobenthic communities (primer sequences and references are displayed in Table 3.2). For the mlCOIintF/LoboR1 primer, PCR reactions were performed for each sample using KAPA HIFI HotStart PCR Kit according to manufacturer instructions, 0.3 µM of each PCR primer and 50 ng of template DNA in a total volume of 25 µL. For the second COI primer-pair (LCO1490/III_C_R) PCR reactions were performed using 1x Advantage 2 Polymerase Mix (Clontech, Mountain View, CA, USA), 0.2 µM of each PCR primer and 25 ng of template DNA in a total volume of 25 µL. For the 18S V4 region, PCR reactions were performed for each sample using KAPA HIFI HotStart PCR Kit according to manufacturer instructions, 0.3 µM of each PCR primer and 12.5 ng of template DNA in a total volume of 25 µL. The production of amplicon libraries and the high-throughput sequencing (HTS) were carried out at Genoinseq (Cantanhede, Portugal). Negative controls for DNA extraction and PCR, and positive PCR controls were included in the molecular analyses' workflow.

Second PCR reactions added indexes and sequencing adapters to both ends of the amplified target region according to manufacturer's recommendations (Illumina, 2013). PCR products were one-step purified and normalized, pooled and pair-end sequenced in an Illumina MiSeq® platform according to manufacturer's instructions (Illumina, San Diego, CA, USA).

The granite substrate of mobile fauna from 3 months of deployment did not amplify with mlCOIintF/LoboR1 primer-pair and was not included in further analyses.

HTS reads were processed as in Chapter 3 (see 3.3.2.1 on Materials and Methods section for more details). The usable reads were then submitted to mBrave (www.mbrave.net, Ratnasingham, 2019) and SILVAngs database (https://ngs.arb-silva.de/silvangs/, Quast et al., 2013), to generate the operational taxonomic unit (OTU) tables and taxonomic assignments for COI and 18S data sets, respectively. In both cases, species-level taxonomic assignments were attributed when displaying ≥97% similarity with reference sequences. Only reads with match at species-level were used for further analysis, and singletons, rare sequences (i.e. less than 8 sequences) and reads matching non-metazoan were discarded (following Lobo et al., 2017). The validity of the species names was verified in World Register of Marine Species (WoRMS) database.

4.2.5 Community statistical analyses

The proportion of overlapping and unique species between substrates and sampling times was determined for both approaches (morphology and DNA metabarcoding) and displayed using Veen diagrams, obtained with the R package VennDiagram (Chen and Boutros, 2011); while qualitative distribution of species among phyla was displayed through barplots (GraphPad Software, Inc.).

Community analysis were performed using Primer v6.1.16 software (Primer-E Ltd, Plymouth, UK; Clarke and Gorley, 2006). Since we had quantitative data from morphological identifications, species richness (*S*), Shannon-Wiener diversity index (*H'*, $\log_2$), Margalef species richness index (*d*) and Pielou's evenness index (*J'*) were calculated for each morphological replicate (i.e. substrate/sampling time) individually. To test possible differences associated to the microhabitat of each substrate replicate, a two-way ANOVA was previously performed (GraphPad Software, Inc.), and no significant differences were found ($p>0.05$).

Multivariate analyses were carried out considering presence/absence of the taxa due to the qualitative nature of the molecular data. Bray-Curtis measure of similarity for presence/absence of species was used to compare morphological identification (pooled data from 3 replicates) and metabarcoding data, and to investigate differences between the four sampling times. The hierarchical clustering (CLUSTER; linkage method: UPGMA) was performed to investigate the coastal macrozoobenthic community similarity among substrates and sampling times between both species-identification methodologies. Non-metric multidimensional scaling (nMDS) analyses were performed based on Bray-Curtis resemblance coefficient to visualize community distribution from distinct methodologies (morphology and DNA metabarcoding) and for all substrates among sampling times (PAST v4.03; Hammer et al., 2001). One-way analyses of similarities (ANOSIM) were used to test for differences in the colonized community between the three substrates. The Similarity Percentages Test (SIMPER) was used to test for differences in the composition of the macrozoobenthic communities in the three types of substrates and among sampling times, between both species-identification methodologies, and to identify the taxa which most accounted for the similarities within each substrate and the differences between them (one way, Bray-Curtis similarity, 70% cut-off). A multivariate ANOVA based on dissimilarities (PERMANOVA v1.0.6) was tested (method: Bray-Curtis, number of permutations: 1000), to obtain the effects of substrates (three factors: Slate, PVC and Granite) and sampling times (four factors: 3, 7, 10 and 15 months) on the structure of macrozoobenthic communities. We compiled the Linnean taxonomic classification (species, genus, family, order, class and phylum) for the species detected in all substrates and sampling times to calculate the average pair-wise

path lengths to estimate assemblages' taxonomic diversity. For the abundance data obtained through morphological identifications, we estimate the Average Taxonomic Distinctiveness (AvTD - Δ^*; Warwick and Clarke, 1995) and the Taxonomic Diversity (TD - Δ) for all replicates of each substrate among sampling times. For presence/absence data we only estimate AvTD (Δ^+; Clarke and Warwick, 1998) for each substrate among sampling times.

Species occurrence was assessed for the consideration of their exclusivity (i.e. species only detected in one substrate/sampling time combination), partial exclusivity (i.e. species present in one sampling time) and ubiquity (i.e. species shared among all substrates and sampling times).

A heatmap was built with the R package "gplots" (Warnes et al., 2020) using the presence/absence data in the different substrates and sampling times. To simplify this analysis, exclusive species (Table S.4.1) and ubiquitous species (*Perforatus perforatus, Mytilus galloprovincialis, Harmothoe impar, Spirobranchus triqueter, Asterocarpa humilis*) were not included.

4.3 Results

Different species corresponding to marine macrozoobenthic taxa colonized all three types of artificial substrates deployed in Toralla Island, and were registered during the experiment. Analyzing the photos take from each substrate in all sampling times (Fig. 4.2), were perceptible differences between sampling times, mostly after 7 months of deployment. However, within sampling times, substrates displayed very similar patterns of colonization.

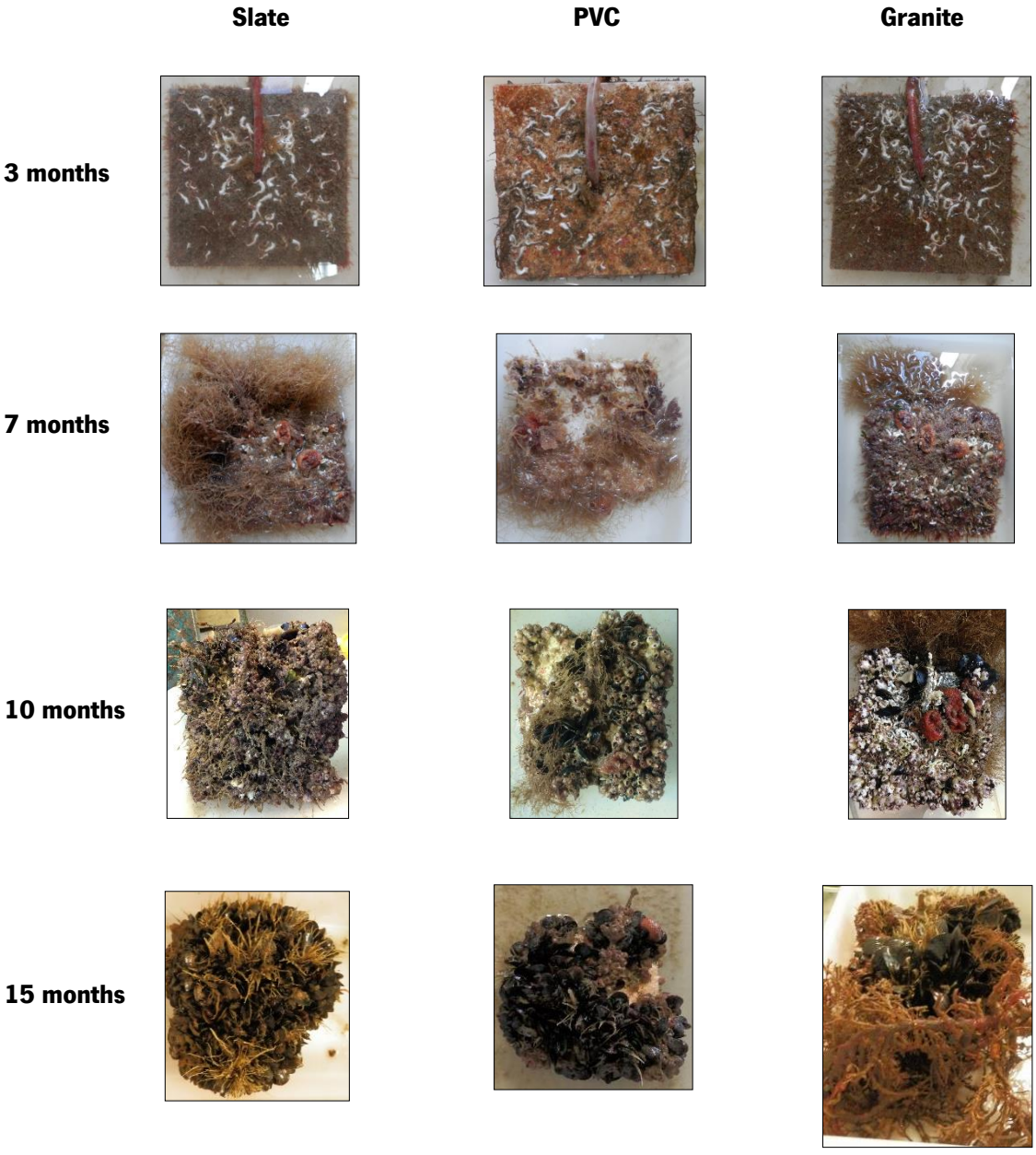


Figure 4.2. Photos of the artificial substrates (slate, PVC and granite) collected after 3, 7, 10 and 15 months of deployment.

4.3.1 Morphology-based macrozoobenthic taxonomic assignments

In total, 15,525 individuals for a total of 71 different taxa, corresponding to 7 zoological phyla (Annelida, Bryozoa, Crustacea, Echinodermata, Hydrozoa, Mollusca and Tunicata), were identified through morphology (Fig. S.4.2). The taxonomic classification of all taxa detected is available in Table S.4.2.

The calculation of univariate indices revealed a total abundance (*N*) from 61 to 1,562 individuals, with an average value of 472 ± 340 individuals/sample (i.e. individuals detected in each substrate per sampling time and replicate), and the macrozoobenthic species richness (*S*) varied between 5 and 33 species, with an average of 18 ± 8. The Margalef species richness index of diversity (*d*) ranged from 0.68 to 5.02, with an average of 2.87 ± 1.14. The *J'* index ranged from 0.10 to 0.90 with an average value of 0.61 ± 0.16 and *H'*($\log_2$) index ranged from 0.16 to 2.74, with an average of 1.75 ± 0.61 (Table 4.1). Although species richness and Shannon-Wiener index did not revealed statistical differences between substrates, statistical differences were detected using species abundance and Magalef index (Two-way ANOVA, $p<0.001$).

Table 4.1. Diversity analysis for the univariate indices of the assemblages detected in each substrate type across sampling times, using morphology. *S* – species richness. *N* – species abundance. *d* – Margalef species richness index. *J'* – Pielou's evenness index. *H'* – Shannon-Wiener index. 3M – 3 months; 7M – 7 months; 10M – 10 months; 15M – 15 months. 1, 2 and 3 corresponds to the replicate number.

Sampling time	Substrate	*S*	*N*	*d*	*J'*	*H'*(log$_2$)
3M	Slate 1	13	186	2,296	0,6947	1,782
	Slate 2	13	144	2,415	0,5998	1,539
	Slate 3	13	164	2,353	0,5016	1,287
	PVC 1	13	92	2,654	0,7762	1,991
	PVC 2	8	61	1,703	0,8996	1,871
	PVC 3	8	110	1,489	0,6391	1,329
	Granite 1	8	142	1,412	0,5177	1,077
	Granite 2	12	222	2,036	0,613	1,523
	Granite 3	10	229	1,656	0,6932	1,596
7M	Slate 1	31	745	4,536	0,7999	2,747
	Slate 2	32	643	4,794	0,7539	2,613
	Slate 3	33	585	5,022	0,7699	2,692
	PVC 1	26	762	3,767	0,5859	1,909
	PVC 2	27	605	4,059	0,6259	2,063
	PVC 3	21	631	3,102	0,6042	1,839
	Granite 1	23	448	3,604	0,6587	2,065
	Granite 2	24	1168	3,256	0,6239	1,983
	Granite 3	21	688	3,061	0,6282	1,913
10M	Slate 1	25	471	3,899	0,5413	1,742
	Slate 2	29	829	4,167	0,6204	2,089
	Slate 3	26	440	4,107	0,7218	2,352
	PVC 1	21	734	3,031	0,7672	2,336
	PVC 2	20	311	3,31	0,5622	1,684
	PVC 3	19	252	3,255	0,6526	1,921
	Granite 1	25	1562	3,264	0,6338	2,04
	Granite 2	25	1051	3,45	0,7795	2,509
	Granite 3	25	354	4,089	0,7565	2,435
15M	Slate 1	9	317	1,389	0,4313	0,9476
	Slate 2	10	426	1,487	0,3056	0,7037
	PVC 1	9	275	1,424	0,4313	0,9477
	PVC 2	5	372	0,6758	9,83E-02	0,1581
	PVC 3	12	336	1,891	0,365	0,907
	Granite 1	12	209	2,059	0,528	1,312

For all substrates and among all sampling times, Crustacea, particularly Amphipoda, was the dominant taxonomic group in both species and abundance (Slate: 34 species; PVC: 30 species; Granite: 28 species), followed by Mollusca, especially Gastropoda, (Slate: 15 species; PVC: 12 species; Granite: 13 species) (Fig. 4.3).

Overall, after 7 months of deployment (July) was identified the highest number of species and individuals, 52 and 6,275 respectively. While the sampling time with the lowest number of species was after 15 months of deployment (March 2018), the lowest number of individuals was verified after 3 months of deployment (March 2017). The substrate with higher number of specimens was granite (6,073 individuals), while slate and PVC had similar number of identified individuals: 4,950 and 4,541, respectively. The number of species detected was similar between substrates, although slate had more species (62 in total) comparatively to others (vs 53 in PVC and 52 in Granite). Statistical differences between sampling times were detected (Two-way ANOVA, $p<0.001$).

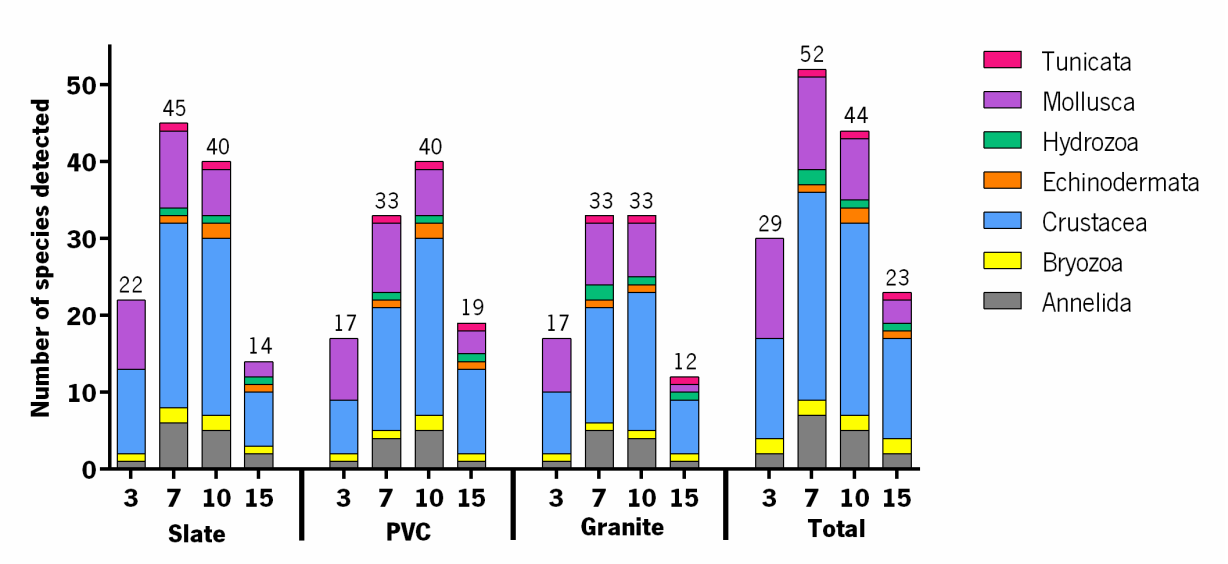


Figure 4.3. Number of detected taxa using morphology in each substrate type and sampling time (3, 7, 10 and 15 months). Total: total number of detected taxa in all substrates combined for each sampling time.

The taxonomic diversity and AvTD for abundance data from each replicate of substrates among sampling times varied between samples and revealed that most of the samples had a lower degree of phylogenetic diversity (TD = 78.90 ± 9 and AvTD = 89.66 ± 7; Table S.4.3). The highest value of AvTD (combined the three replicates) was detected for slate deployed for 15 months, and the lowest for granite after 10 months of deployment. Furthermore, we calculate the relation between AvTD and 3, 7 and 10 sampling times (we

exclude 15 months from this analysis because we do not have three replicates of each substrate), and results showed that for most substrates decreased with sampling time (Fig. S.4.3 A). These results revealed a different tendency from species richness for each substrate between sampling times (Fig. S.4.3 B).

4.3.2 DNA metabarcoding-based macrozoobenthic taxonomic assignments

Using the three primer-pairs amplifying the barcode regions of two marker loci (COI-5P and 18S V4 region), a total of 171 different taxa distributed along 9 different phyla (Annelida, Bryozoa, Crustacea, Echinodermata, Hydrozoa, Mollusca, Nemertea, Platyhelminthes and Tunicata) were identified on the three artificial substrates (Fig. 4.4). The detailed taxonomy of the detected taxa is available in Table S.4.2.

The highest species richness was detected after 7 months of deployment, in the three substrates. Globally, the three substrates had a similar number of species detected: 124 species detected in slate and 123 species detected in both PVC and granite. The three substrates were dominated by the three major groups of macrozoobenthos: Mollusca, Annelida and Crustacea (70.2% of the total species for slate and granite and 67.5% of the total species detected for PVC). Statistical differences between sampling times were detected (Two-way ANOVA, $p<0.001$).

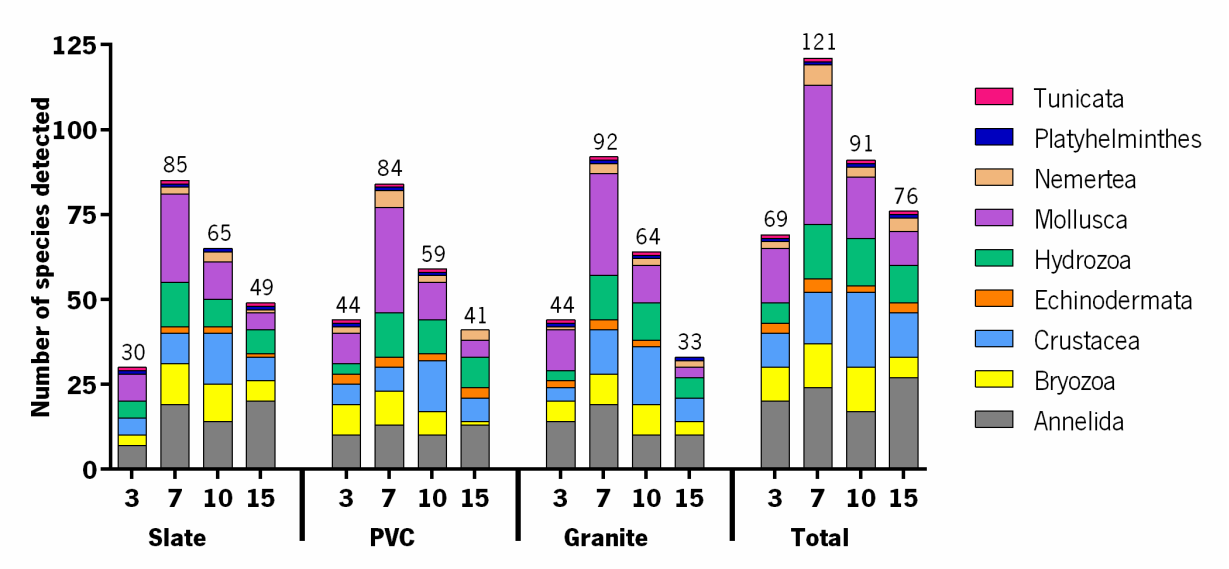


Figure 4.4. Number of detected taxa using DNA metabarcoding in each substrate type and sampling time (3, 7, 10 and 15 months). Total: total number of detected taxa in all substrates combined for each sampling time.

4.3.3 Morphology and DNA metabarcoding comparison

The taxonomic profile of the communities was highly dependent on the identification approach adopted (Fig. 4.5). Globally 203 different taxa were recorded. Compared to morphology, DNA metabarcoding retrieved more species in all substrate/sampling time combinations, namely when comparing with the pooled data for the morphological replicates. Additionally, a higher number of phyla was also detected using DNA metabarcoding: Platyhelminthes and Nemertea were not recorded using morphology. The community detected in each substrate/sampling time combination was significantly different ($p<0.0001$) when using morphology or DNA metabarcoding approaches.

A high proportion of taxa could not be identified to species using morphology, contributing to the lower number of species detected compared to DNA-based approaches. However, 32 species identified with morphology were not detected using DNA metabarcoding (e.g. the polychaete *Lepidonotus squamatus*, or the caprelid *Pseudoprotella phasma*), and both approaches demonstrated to be complementary in their ability to detect marine macrozoobenthic species (Fig. 4.5). From the total species recorded, 15.8% were exclusively detected through morphology and 65.0% through DNA metabarcoding, while only 19.2% were detected by both approaches. The dominant taxonomic groups were different for both approaches: while in morphology Crustacea was the dominant group (39 species detected in total), Mollusca was the taxonomic group with more representative species through DNA metabarcoding (48 species detected in total). Annelida yielded over five times more species detected with DNA metabarcoding (7 species detected in total with morphology vs 33 with DNA metabarcoding). Comparing the substrates in both approaches, the three substrates have similar patterns on species detection, and the highest number of species was recovered after 7 months.

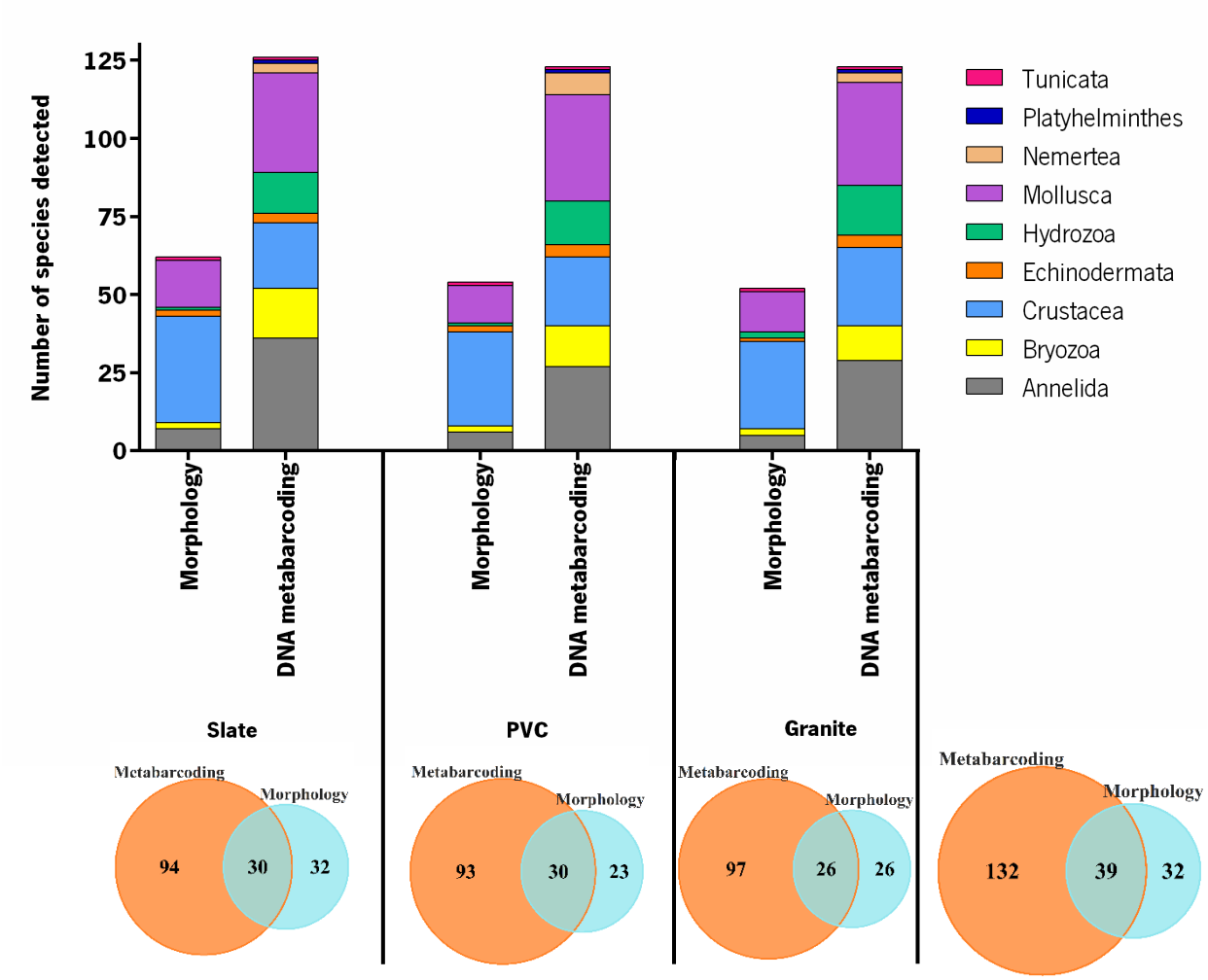


Figure 4.5. Comparison between morphological and DNA metabarcoding species-level identifications in three substrates types (Slate, PVC and Granite) with sampling time data merged. The upper bar chart shows the distribution of the total number of species per phylum obtained either by morphology or DNA metabarcoding, in each substrate type. The Venn diagrams in the lower part shows the proportion of species detected exclusively by morphology (blue), exclusively by DNA metabarcoding (orange), and shared by both approaches (overlapping circles), for each substrate type and for all combined substrates. Differences were detected between both approaches (cophenetic correlation coefficient of the Bray-Curtis similarity = 0.51).

Non-metric multidimensional analysis (Fig. 4.6), comparing the species detected by both identification methodologies, showed two distinct groups: all DNA metabarcoding samples clustered together and separated from the samples identified through morphology. Furthermore, except the slate substrate collected after 3 months, the combinations substrate/sampling time clustered by sampling time. SIMPER analysis revealed high dissimilarity between the species assemblages detected by both approaches (average dissimilarity of 76.97%).

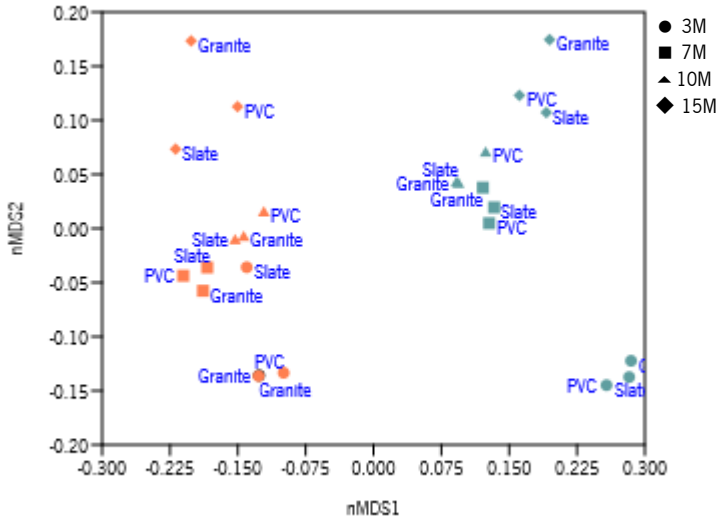


Figure 4.6. Non-metric multidimensional scaling (nMDS), based on Bray-Curtis similarity index, of marine macrozoobenthic species detected in the artificial substrates (Slate, PVC and Granite) deployed in Toralla Island. Colors group species identified by different methodologies: morphology (blue) and DNA metabarcoding (orange). Sampling times: 3M – 3 months; 7M – 7 months; 10M – 10 months; 15M – 15 months.

4.3.4. Ecological successional patterns: changes in community composition

The composition of communities slightly differs between substrates (Fig. 4.7 and S.4.4), and the detection of specific species seems to be dependent on the used substrate (Fig. 4.8). We detected a higher occurrence of exclusive species (approximately 27%) than partially exclusive species (approximately 5%). Furthermore, ubiquity was only detected for 5 species (approximately 2%), indicating a variable occurrence of species among sampling times, regardless of the substrate.

Although the contribution of each taxa was similar between the three substrates, the total number of species was slightly different among substrates. Substrates were largely dominated by mollusks (56

species), crustaceans (53 species) and annelids (47 species). Compared to the other substrates, slate replicates resulted in a slightly higher rate of species detection (154 species) and PVC replicates had the lowest diversity (142 species). Statistical differences between sampling times were detected (Two-way ANOVA, $p<0.001$).

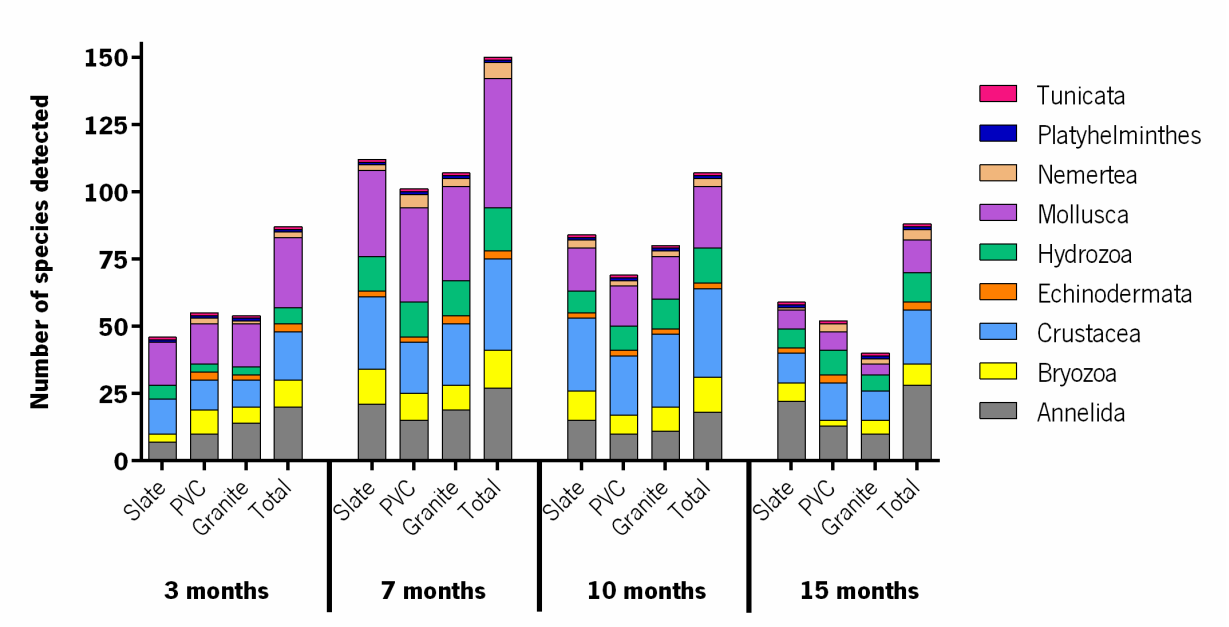


Figure 4.7. Total detected taxa for each artificial substrate (slate, PVC and granite) and for the total species detected between the three substrates (Total), among sampling times (3, 7, 10 and 15 months).

Almost all of the species was detected in the three substrates (47.8%) and only a small percentage was exclusive to a particular substrate (11.8% in slate to 7.4% in PVC; Fig. 4.8). Between each sampling time, the number of shared species between substrates was higher, which only a small percentage of species were only detected by one substrate. However, the very exclusive species occurred only in one sampling time and are probably rare species.

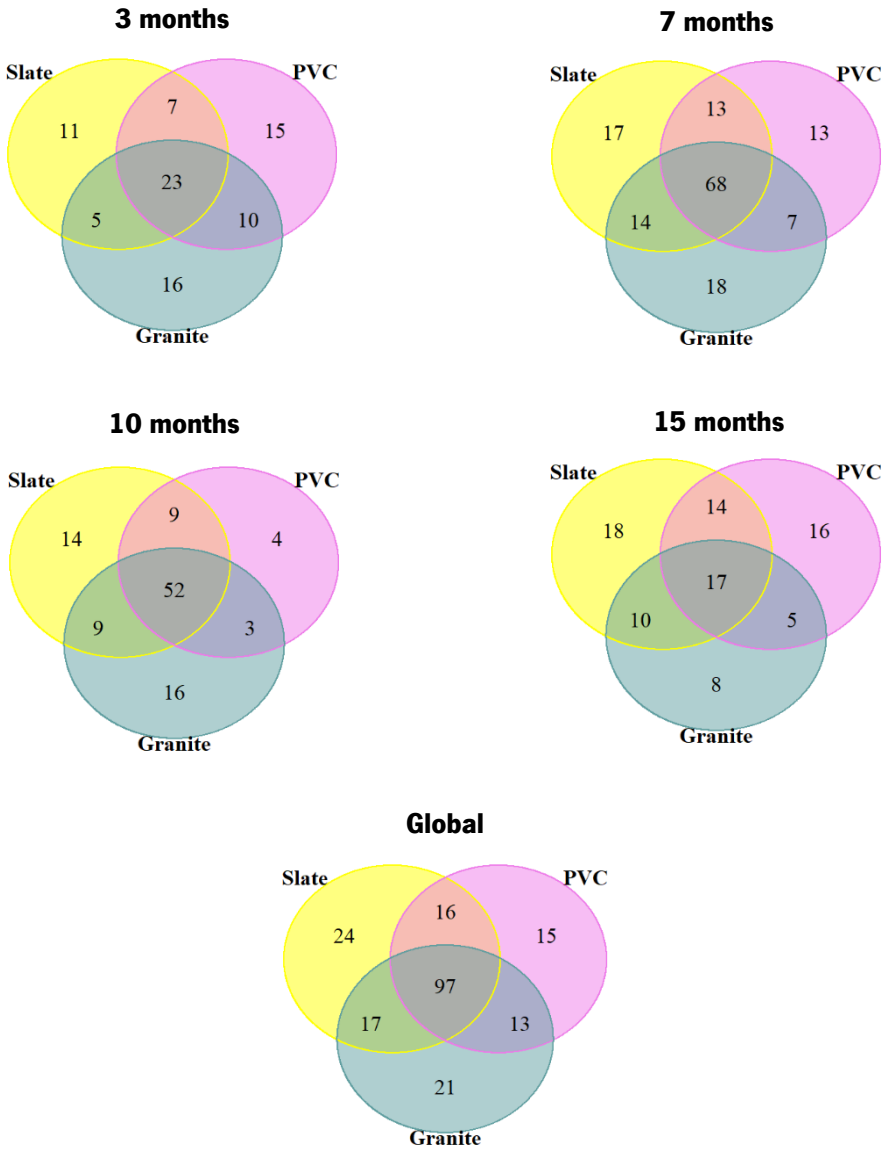


Figure 4.8. Proportion of the species detected exclusively and shared between the three substrates (Slate, PVC and Granite) for each sampling time (3, 7, 10 and 15 months), and for the combination of the total species detected in each substrate (global).

The heatmap reveals detailed differences in the occurrence of species among the four sampling times and the three artificial substrates (Fig. 4.9). The largest increase of species detection was observed from 3 to 7 months of deployment (more 63 species detected), and on the following sampling times a decrease in species recovery was recorded, in total and for each substrate (Fig. 4.7).

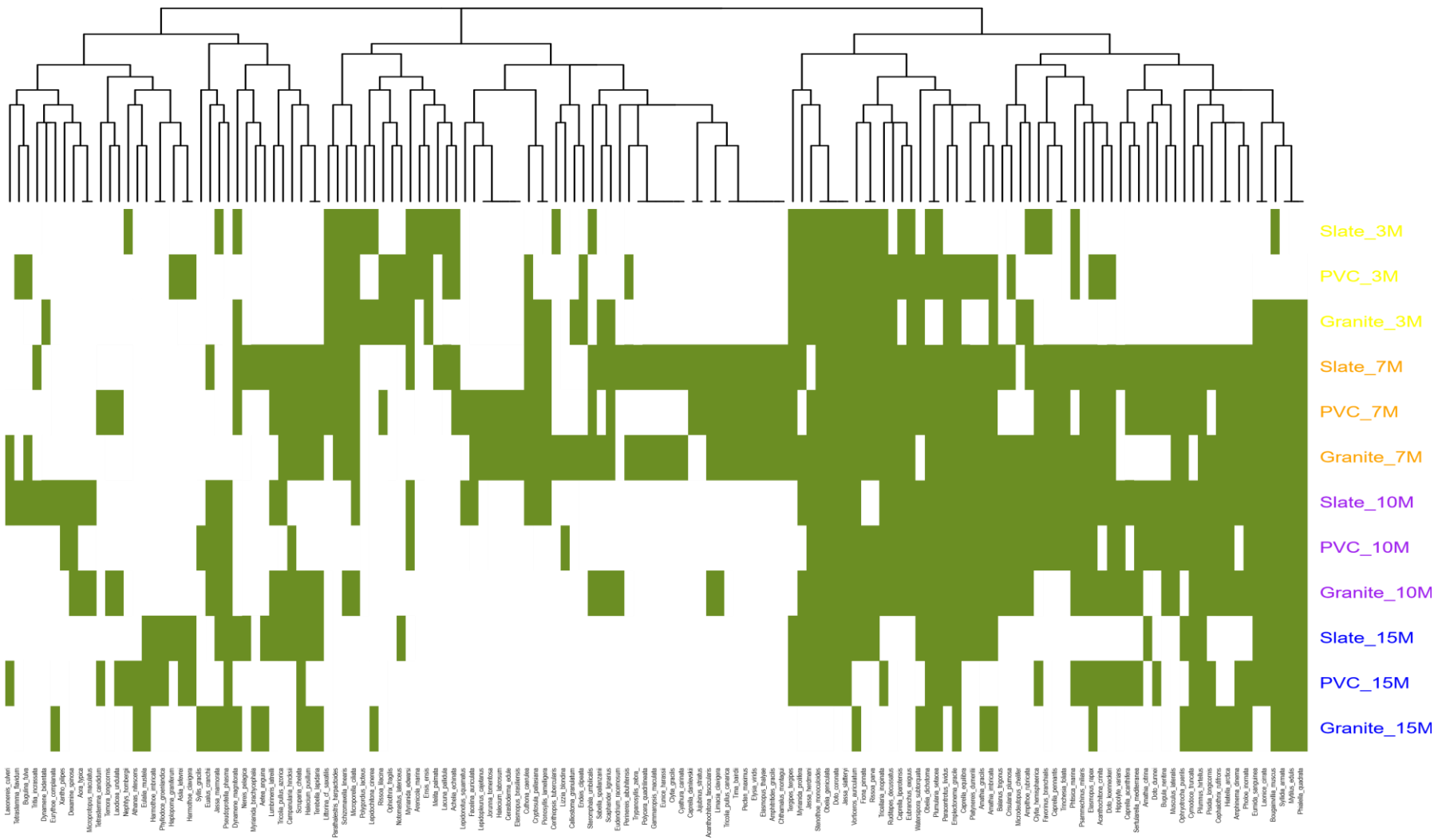


Figure 4.9. Heatmap of marine macrozoobenthic species detected in each substrate and sampling time. Substrates are grouped by colors corresponding to months of deployment; 3M – 3 months; 7M – 7 months; 10M – 10 months; 15M – 15 months. Green represents species occurrence.

In a SIMPER analysis, after 10 months of deployment was detected the greatest similarity between species composition for all substrates combined. On the other hand, between 3/15 sampling times had the greatest compositional shifts, with an average dissimilarity of 61.82% (Table 4.2). Furthermore, between 7 and 10 months of deployment the species composition were more similar than compared with other sampling times. Crustacea and Mollusca, two of the most representative taxa, were the main drivers of the differences between sampling times (Table S.4.4).

Table 4.2. SIMPER analysis showing species-level differences in community composition across all sampling times. 3M – 3 months; 7M – 7 months; 10M – 10 months; 15M – 15 months.

Sampling time	Average similarity within sampling	Sampling times comparison	Average dissimilarity between sampling
		3M/7M	54.49
3M	58.65	3M/10M	53.41
7M	74.35	3M/15M	61.82
10M	76.33	7M/10M	35.84
15M	52.74	7M/15M	55.73
		10M/15M	46.62

The AvTD funnel plot (AvTD = 92.9 ± 0.7; Fig. 4.10 and Table S.4.5) showed that most of the substrates/sampling time combinations are lower than the 95% confidence interval. Although none sample is lower than funnel limits, differences suggest a lower degree of taxonomic diversity.

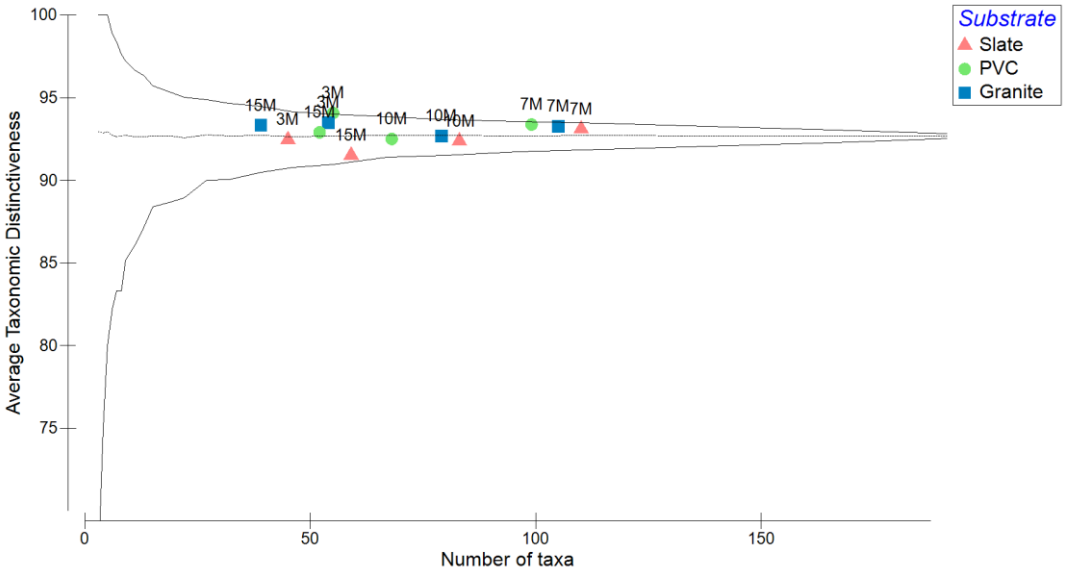


Figure 4.10. Average taxonomic distinctiveness (AvTD) funnel plot of the marine macrozoobenthic communities recorded in artificial substrates (slate, PVC and granite) among sampling times. 3M – 3 months; 7M – 7 months; 10M – 10 months; 15M – 15 months.

Cluster analysis (Fig. S.4.5) and non-metric multidimensional scaling (Fig. 4.11) comparing the species detected in each sampling time revealed different time-patterns of taxa to colonize the substrates: substrates from the same sampling time cluster together preferentially, regardless of the substrate type. Ten species were detected in all substrates but only in one sampling time (e.g. the gastropod *Elysia viridis*, detected for all substrates after 7 months of deployment). However, most of the species detected had a low occurrence (between substrates and sampling times), and 95 species (46.8%) were exclusively recovered in only one or two combination substrate/sampling time. Other species were randomly detected, showing variations in occurrence among substrates and sampling times (e.g. the bryozoan *Bougainvillia muscus* or the amphipod *Jassa slatteryi*). On the other hand, 5 of the species were recorded in all substrates and sampling times (*Asterocarpa humilis, Harmothoe impar, Mytilus galloprovincialis, Perforatus perforatus and Spirobranchus triqueter*). The most common species belonged to Crustacea, followed by Annelida and Mollusca.

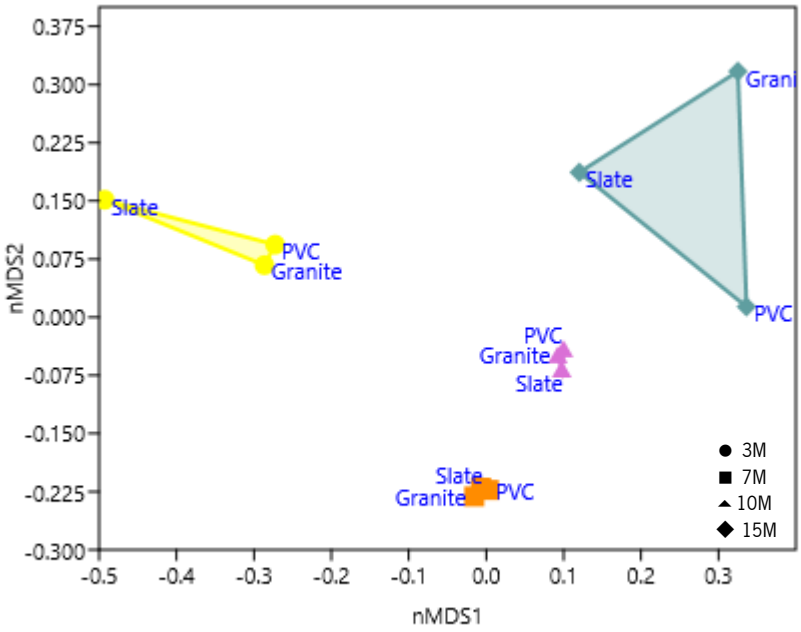


Figure 4.11. Non-metric multidimensional scaling (nMDS), based on Bray-Curtis similarity index, of the total species detected in the artificial substrates (Slate, PVC and Granite) among sampling times. Sampling times: 3M – 3 months (yellow); 7M – 7 months (orange); 10M – 10 months (purple); 15M – 15 months (blue).

4.4 Discussion

This study consisted on a comprehensive time-series assessment of the use of different artificial substrates, combined with morphology and DNA-based taxonomic profiling, for monitoring coastal macrozoobenthic assemblages. To our best knowledge, the employment of two genetic markers and three sets of primer-pairs in testing DNA-based species detection in these target communities is unprecedented. The 12 different substrate-sampling time combinations provide plenty of opportunity to compare the performance of morphology and DNA-based species detection. Clearly the DNA-based approach had consistently much higher detection ability, but there was still a sizeable proportion of species detected exclusively using morphology.

Globally, the substrate type did not show a significant impact on the taxonomic composition of the assemblages, except on the initial sampling times. However, species composition differed significantly between sampling times, providing insights into colonization timings and ecological succession patterns in these communities, as well as the impact of these factors for artificial substrate-based monitoring. These findings and their implications are discussed in detail below.

4.4.1 Comparison of species detection success employing morphology or DNA-based approaches

The diversity of species detected through DNA metabarcoding and morphology were dominated by the same taxa (Annelida, Crustacea and Mollusca). However, compared to morphology DNA metabarcoding retrieved more species. The results also revealed underestimation of some taxonomic groups depending on used approaches (e.g. hydrozoans or bryozoans using morphology), resulting in a different taxonomic composition provided by each identification methodology. These results are consistent with several studies, which demonstrated that DNA metabarcoding can capture more diversity than morphological identification for different taxa: estuarine macrobenthos (Lobo et al., 2017); freshwater macroinvertebrates (Elbrecht et al., 2017); non-indigenous species (von Ammon et al., 2018); meiofaunal eukaryotes (Haenel et al., 2017); diet of marine fishes (Berry et al., 2015); seagrass-associated communities (Cowart et al., 2015); zooplankton (Schroeder et al., 2020).

In morphology-based identifications, we found no clear relationships between taxonomic indices (AvTD) and species richness (*S*) in the three replicates of each substrate among sampling times. The lower degree of phylogenetic diversity detected for the majority of samples, largely diverge from the highest species richness obtained after 7 and 10 months of deployment. Interestingly, the replicates with more species richness were not the one with highest phylogenetic diversity. For example, for slate replicates after 7 months of deployment, we detected 45 species, the highest species richness for that substrates, and lower values of taxonomic distinctiveness (89.98%). These results revealed that assemblages with higher species richness do not necessarily mean a higher phylogenetic diversity, than others with less species.

The lower capacity of morphology for species-level identifications was likely due to the occurrence of larvae and other early stages of development, as well as to the presence of very damaged or fragmented specimens due to the sieving process (e.g. damaged specimens of polychaetes). For example, a number of decapod specimens at post-larvae stage were identified through morphology only as megalopa. Some other taxonomically difficult groups, such as the *Eumida* species complex (Teixeira et al., 2019), limited the species-level taxonomic assignments using morphology. These results were expected, since previous studies already demonstrated higher taxonomic discrimination using DNA metabarcoding (Berry et al., 2015; Gibson et al., 2015; Serrana et al., 2019; Porter and Hajibabei, 2020). This highlights the strengths of DNA metabarcoding in generating more reliable and accurate identifications in marine macrobenthos monitoring.

Although DNA metabarcoding demonstrated to be more efficient in detecting marine macrozoobenthic species, there were still 16% of species that were identified morphologically but not with DNA metabarcoding. However, five of them still lack DNA barcode in both used databases (e.g. the amphipod *Elasmopus thalyae*). In Chapter 3 (section 3.4.3), the analysis on reference sequences in used databases (BOLD and SILVA) already revealed a gap of DNA barcodes for the species detected, and the incompleteness and inaccuracy of used databases are currently a recognized problem (Leite et al., 2020). These explain the impossibility to detect that five species using DNA metabarcoding, and highlight the necessity to complete reference databases, especially with species from the sampling area.

Even with reference sequences available on databases, other 27 species were not detected by DNA metabarcoding. Some of these species are exclusive species (i.e. species detected only in one sample) or have been reported as cryptic complexes (e.g. the annelid *Trypanosyllis zebra*). In the latter case, the possibility that they correspond to cryptic lineages whose sequences are not available yet in genetic databases cannot be discarded. These failures on DNA-based species detection may be ascribed to possible low amounts of DNA available for amplification, in the case of small-size / low abundance specimens, together with insufficient sequencing depth. Actually, higher sequencing depth may be required to capture exclusive species (Porter and Hajibabaei, 2018; McGee et al., 2019). Alternatively, it could be a result of the performance of molecular methodological steps, such as DNA extraction or primer-pair affinity's problems (Piñol et al., 2015; Elbrecht et al., 2017; Serrana et al., 2019). Furthermore, since the replicate samples used for morphology were not the same as the ones used for DNA metabarcoding, the non-detected species could simply be missing in the replicate used for the DNA-based approach. This result led to recommend a combined analysis, where morphology will be used in a higher time-scale strategy and DNA-based identification approaches more frequently (time and space), to confirm DNA metabarcoding results and to obtain species abundances.

Although using morphology we detected *Mytilus galloprovincialis*, a very representative mollusk from Galician coast and with high representativeness of specimens in the substrates, with DNA metabarcoding (for both COI and 18S markers) it was not possible to detect that mussel at species-level. This is a result of inability of molecular markers to identify that species, probably due to a common problem of taxonomic uncertainties of this genus (Śmietanka et al., 2004) combined with databases errors and misidentifications. Databases are replete of ambiguous sequences, where this non-accurate data for *Mytillus* spp. normally result in an identification attributed to different *Mytilus* species. Some of these ambiguous sequences are

a product of misidentifications, resulting in unreliable species identifications. Furthermore, molecular taxonomic uncertainty is commonly associated with these species due to hybridization and divergence between male and female mtDNA (Śmietanka et al., 2017). Additionally, problems associated to extraction procedures or primers efficiency could also influence the ability to detect that species. Cahill and collaborators (2018), found similar unrecovery problems with *M. galloprovincialis* using DNA metabarcoding and associated them to poor mismatches to the forward primer, which was the same that we used: mlCOIintF. Other study using mock communities, associated the variations in the ability of primer-pair to detect species (especially for *Mytilus* sp.) to non-specificity of the primers and random sampling during PCR or sequencing (Hollatz et al., 2016). Although we cannot obtain a DNA-based robust species identification, we conclude that we probably detected the same species using both approaches, *M. galloprovincialis*. These results also highlight the need to reliable and accurately complete reference databases.

4.4.2 Substrates effects and seasonal succession patterns on macrozoobenthic community

Compared to the other substrates, slate resulted in a higher rate of species detection. However, any substrate was able to detect all the species. Actually, a difference in the number of species detected between substrates was observed, and exclusive species were recorded in all substrates. Although the deployment method was the same for all substrates, different weight of plates led to flotation of PVC plates which could affect the actual depth and consequently the composition of the community (Glasby, 2001). Furthermore, changes in hydrodynamics and irradiance times on the surfaces of the substrates influence the settlement of species on the substrates (Field et al., 2007). These results highlight that using a single and structurally simpler substrate will result in less ability to detect the diversity of macrozoobenthic taxa from a marine community. We propose a strategy using different substrates with tridimensional structure, which will create more habitat complexity and different area for settlement and refuge, and allow to achieve the taxonomic diversity of a full community.

Consistent with the results obtained for morphological data, combining the results obtained with morphology and DNA metabarcoding, the AvTD results for the species detected in each substrate among four sampling times revealed lower levels of taxonomic diversity (<95% confidence limit). However, the highest number of species was detected between 7 and 10 months of deployment, the sampling time with highest phylogenetic diversity was after 3 months. Furthermore, compared to 10 months of deployment,

at 15 months we detected less species in the PVC and granite, however with higher taxonomic diversity levels. These results revealed a similar degree of complexity in the structure of the communities detected between substrates, and a low relationship between the number of species detected and phylogenetic diversity.

Only in PVC substrates we detected spaces without covered. However, clear spaces in substrates without colonization should not be associated to the number of species detected. Other studies have suggested a non-correlation between surface structure and colonization (Antoniadou et al., 2010). From the used substrates in this study, PVC is one of the most commonly applied for species colonization, however comparing results from other studies using artificial substrates is challenging. The lack of standardization between studies, based on many types and size of substrates as well as the method of deployment can influence species recruitment and the outcome of succession of macrozoobenthic species. Contrary to our results, difficulties in primary colonization of mobile organisms in PVC plates were detected in previous studies (Adey and Vassar, 1975). In a study using artificial substrates to investigate successional patterns of macrobenthic communities, authors obtained a higher colonization ability associated to more structured and heterogeneous surfaces (Uribe et al., 2015).

The highest diversity of species was attained after 7 months of deployment for all substrates. Sampling efforts and processing can explain different species detection ability depending on substrate type. For example, the clam *Ruditapes decussatus,* a typical sediment-associated organism, was detected by both approaches however with variations between substrate types along sampling times. These results suggest that habitat complexity increased during the process of colonization, and consequently more food, shadow and refuge areas were available for diverse faunal settlement.

The growth and development of the community varied according to the selected substrate, evidencing that the combination between sampling times and substrate determined the taxonomic composition of the community: between each sampling time new species were recovered and others ceased to be recorded. This stochastic colonization may be associated with the seasonality, namely the changes commonly associated with season (e.g. recruitment season, presence of predators or temperature) which influence the resultant community. For example, in a previous study developed in the same study area, mussels dominated the community after 6 months of deployment (Leal et al., 1994), however in our study this happened only after 15 months of deployment. Furthermore, the differences detected between 3 and 15 months of deployment (i.e. highest dissimilarity detected between these sampling times) suggested

variations in the structure and complexity of the habitat. This is also supported by the decrease in the number of species detected after 7 months, which is probably associated to the competition for space or migration of some species. Succession studies using artificial substrates also demonstrated differences in communities as different species in advance of sampling times were added or lost (Antoniadou et al., 2010; Pacheco et al., 2010; Uribe et al., 2015).

4.5 Conclusion

Artificial substrates promote macrozoobenthic colonization and coupled with molecular tools provides the ability for a high-throughput strategy for monitoring coastal macrozoobenthic communities. Through DNA metabarcoding we obtained a much deeper diagnosis of species occurrence compared with morphological identifications, highlighting the utility of DNA-based identification approaches for marine macroinvertebrates assessment, providing an improvement toward the development of a reliable, accurate, throughput and cost-effective approach. Although DNA metabarcoding offer advantages over morphology, assigning a much higher number of taxa to species, particularly some taxonomic groups with small size specimens unrecorded using morphology, there was still a fraction of species detected exclusively through morphology. Based on the recognized current limitations of the DNA metabarcoding methodology and the necessity to complete reference libraries, until these issues are solved, where possible both methods should be used to avoid missing relevant taxa. The results also illustrate the influence of season in the recruitment of macrozoobenthos, with no significant differences between the substrates. Initially were detected a stochastic community corresponding to the initial phases of colonization, growing to sequential stages of processes of succession of marine macrozoobenthos, discriminating the community mainly distributed along the three most representative marine phyla. In future studies the choice of artificial substrate should depend on the aim of the study, highlighting the usefulness and non-destructive characteristics of substrates to promote species recruitment and growth for a community settlement. This project provided an opportunity to yield insights for a better comprehension of fundamental ecological processes in macroinvertebrates communities and opening new perspectives to be used in other studies to supply more extensive, detailed and rigorous data for marine resource management and biodiversity conservation in coastal ecosystems.

References

Adey W., Vassar J.M. (1975). Colonization, succession and growth rates of tropical crustose coralline algae (Rhodophyta, Cryptonemiales). Phycologia. 14, 55–69. https://doi.org/10.2216/i0031-8884-14-2-55.1.

Antoniadou C., Voultsiadou E., Chintiroglou C. (2010). Benthic colonization and succession on temperate sublittoral rocky cliffs. Journal of Experimental Marine Biology and Ecology. 382, 145–153. https://doi.org/10.1016/j.jembe.2009.11.004.

Aylagas E., Borja Á., Muxika I., Rodríguez-Ezpeleta N. (2018). Adapting metabarcoding-based benthic biomonitoring into routine marine ecological status assessment networks. Ecological Indicators. 95, 194-202. https://doi.org/10.1016/j.ecolind.2018.07.044.

Berry O., Bulman C., Bunce M., Coghlan M., Murray D.C., Ward R.D. (2015). Comparison of morphological and DNA metabarcoding analyses of diets in exploited marine fishes. Marine Ecology Progress Series. 540, 167-181. https://doi.org/10.3354/meps11524.

Bowden D.A., Clarke A., Peck L.S., Barnes D.K.A. (2006). Antarctic sessile marine benthos: colonisation and growth on artificial substrata over three years. Marine Ecology Progress Series. 316, 1–16. https://doi.org/10.3354/meps316001.

Cacabelos E., Martins G.M., Thompson R., Prestes A.C.L., Azevedo J.M.N., Neto A.I. (2016). Material type and roughness influence structure of inter-tidal communities on coastal defenses. Marine Ecology. 37, 801–812. https://doi.org/10.1111/maec.12354.

Cacabelos E., Ramalhosa P., Canning-Clode J., Troncoso J.S., Olabarria C., Delgado C., Dobretsov S., Gestoso I. (2020). The role of biofilms developed under different anthropogenic pressure on recruitment of macro-invertebrates. International Journal of Molecular Sciences. 21(6), 2030. https://doi.org/10.3390/ijms21062030.

Cahill A.E., Pearman J.K., Borja A., Carugati L., Carvalho S., Danovaro R., Dashfield S., David R., Féral J., Olenin S., Siaulys A., Somerfield P.J., Trayanova A., Uyarra M.C., Chenuil A. (2018). A comparative analysis of metabarcoding and morphology-based identification of benthic communities across different regional seas. Ecology and Evolution. 8, 8908–8920. https://doi.org/10.1002/ece3.4283.

Cangussu L.C., Altvater L., Haddad M.A., Cabral A.C., Heyse H.L., Rocha R.M. (2010). Substrate type as a selective tool against colonization by non-native sessile invertebrates. Brazilian Journal of Oceanography. 58(3), 219-231. https://doi.org/10.1590/S1679-87592010000300005.

Carew M.E., Kellar C.R., Pettigrove V.J., Hoffmann A.A. (2018). Can high-throughput sequencing detect macroinvertebrate diversity for routine monitoring of an urban river? Ecological Indicators. 85, 440–450. https://doi.org/10.1016/j.ecolind.2017.11.002.

Carvalho S., Aylagas E., Villalobos R., Kattan Y., Berumen M., Pearman J.K. (2019). Beyond the visual: using metabarcoding to characterize the hidden reef cryptobiome. Proceedings of the Royal Society B. 286, 20182697. http://dx.doi.org/10.1098/rspb.2018.2697.

Chen H., Boutros P.C. (2011). VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinformatics. 12, 35. http://dx.doi.org/10.1186/1471-2105-12-35.

Clarke K.R., Gorley R.N. (2006). PRIMER V6: user manual/tutorial. PRIMER-E Ltd, Plymouth.

Clarke K.R., Warwick R.M. (1998). A taxonomic distinctness index and its statistical properties. Journal of Applied Ecology. 35(4), 523-531. https://doi.org/10.1046/j.1365-2664.1998.3540523.x.

Conlan K.E. (1990). Revision of the crustacean amphipod genus *Jassa* Leach (Corophioidea: Ischyroceridae). Canadian Journal of Zoology. 68(10), 2031-2075. https://doi.org/10.1139/z90-288.

Cowart D.A., Pinheiro M., Mouchel O., Maguer M., Grall J., Miné J., Arnaud-Haond S. (2015). Metabarcoding is powerful yet still blind: A comparative analysis of morphological and molecular surveys of seagrass communities. PLoS One. 10(2), e0117562. https://doi.org/10.1371/journal. pone.0117562.

Cristescu M.E. (2014). From barcoding single individuals to metabarcoding biological communities: towards an integrative approach to the study of global biodiversity. Trends in Ecology and Evolution. 29(10), 566-571. https://doi.org/10.1016/j.tree.2014.08.001.

Dowle E.J., Pochon X., Banks C.J., Shearer K., Wood S.A. (2016). Targeted gene enrichment and high-throughput sequencing for environmental biomonitoring: A case study using freshwater macroinvertebrates. Molecular Ecology Resources. 16(5), 1240-1254. https://doi.org/10.1111/1755-0998.12488.

Edgar G.J., Klumpp D.W. (2003). Consistencies over regional scales in assemblages of mobile epifauna associated with natural and artificial plants of different shape. Aquatic Botany. 75(4), 275-291. https://doi.org/10.1016/S0304-3770(02)00194-8.

Elbrecht V., Vamos E.E., Meissner K., Aroviita J., Leese F. (2017). Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. Methods in Ecology and Evolution. 8(10), 1265-1275. https://doi.org/10.1111/2041-210X.12789.

Field S.N., Glassom D., Bythell J. (2007). Effects of artificial settlement plate materials and methods of deployment on the sessile epibenthic community development in a tropical environment. Coral Reefs. 26(2), 279-289. https://doi.org/10.1007/s00338-006-0191-9.

García-Sanz S., Navarro P.G., Tuya F. (2014). Colonization of prosobranch gastropods onto artificial substrates: Seasonal patterns between habitat patches. American Malacological Bulletin. 32, 94–103. http://dx.doi.org/10.4003/006.032.0108.

Gee J.M., Warwick R.M. (1996). A study of global biodiversity patterns in the marine motile fauna of hard substrata. Journal of the Marine Biological Association of the United Kingdom. 76, 177-184. https://doi.org/10.1017/S0025315400029106.

Gibson J.F., Shokralla S., Curry C., Baird D.J., Monk W.A., King I., Hajibabaei M. (2015). Large-scale biomonitoring of remote and threatened ecosystems via high-throughput sequencing. PLoS One. 10, e0138432–15. https://doi.org/10.1371/journal.pone.0138432.

Glasby T.M. (2001). Development of sessile marine assemblages on fixed versus moving substrata. Marine Ecology Progress Series. 215, 37–47. https://doi.org/10.3354/meps215037.

Gouillieux B., Sorbe J.C. (2015). *Elasmopus thalyae* sp. nov. (Crustacea: Amphipoda: Maeridae), a new benthic species from soft and hard bottoms of Arcachon Bay (SE Bay of Biscay). Zootaxa. 3905(1), 107-118. https://doi.org/10.11646/zootaxa.3905.1.6.

Guerra-García J.M., Navarro-Barranco C., Corzo J., Cobos-Munoz V., García-Adiego E.M., Giménez F.S., García-Gómez J.C. (2013). An illustrated key to the soft-bottom caprellids (Crustacea: Amphipoda) of the Iberian Peninsula and remarks to their ecological distribution along the Andalusian coast. Helgoland Marine Research. 67(2), 321-336. https://doi.org/10.1007/s10152-012-0324-1.

Haenel Q., Holovachov O., Jondelius U., Sundberg P., Bourlat S. (2017). NGS-based biodiversity and community structure analysis of meiofaunal eukaryotes in shell sand from Hållö Island, Smögen, and soft mud from Gullmarn Fjord, Sweden. Biodiversity Data Journal. 5, e12731. https://doi.org/10.3897/BDJ.5.e12731.

Hammer Ø., Harper D.A.T., Ryan P.D. (2001). PAST: Paleontological statistics software package for education and data analysis. Palaeontologia Electronica. 4(1), 9.

Hayward P.J., Ryland J.S. (1995). Handbook of the marine fauna of North-west Europe. Oxford University press. Oxford, United Kigdom.

Hebert P.D.N., Cywinska A., Ball S.L., Dewaard J.R. (2003). Biological identifications through DNA barcodes. Proceedings of the Royal Society of London Series B: Biological Sciences. 270(1512), 313-321. https://doi.org/10.1098/rspb.2002.2218.

Hollatz C., Leite B.R., Lobo J., Froufe H., Egas C., Costa F.O. (2016). Priming of a DNA metabarcoding approach for species identification and inventory in marine macrobenthic communities. Genome. 60(3), 260-271. https://doi.org/10.1139/gen-2015-0220.

Illumina. (2013). 16S Metagenomic sequencing library preparation preparing 16S Ribosomal RNA gene amplicons for the Illumina MiSeq system.

Ivanova N.V., de Waard J.R., Hebert P.D.N. (2006). An inexpensive, automation-friendly protocol for recovering high-quality DNA. Molecular Ecology Notes. 6, 998–1002. https://doi.org/10.1111/j.1471-8286.2006.01428.x.

Kelly R.P., Closek C.J., O'Donnell J.L., Kralj J.E., Shelton A.O., Samhouri J.F. (2017). Genetic and manual survey methods yield different and complementary views of an ecosystem. Frontiers in Marine Science. 3, 283. https://doi.org/10.3389/fmars.2016.00283.

Kelly R.P., Port J.A., Yamahara K.M., Crowder L.B. (2014). Using environmental DNA to census marine fishes in a large mesocosm. PLoS One. 9(1), e86175. https://doi.org/10.1371/journal.pone.0086175.

Leal C.B., Vázquez M.I.G. (1994). Efectividad de las pinturas antifouling en el puerto de Vigo (Galicia, NO de la Península Ibérica). NACC: Nova Acta Científica Compostelana, Bioloxía.

Leite B.R., Vieira P.E., Teixeira M.A.L., Lobo-Arteaga J., Hollatz C., Borges L.M.S., Duarte S., Troncoso J.S., Costa F.O. (2020). Gap-analysis and annotated reference library for supporting macroinvertebrate metabarcoding in Atlantic Iberia. Regional Studies in Marine Science. 36, 101307. https://doi.org/10.1016/j.rsma.2020.101307.

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. https://doi.org/10.1073/pnas.1424997112.

Lincoln R. (1979). British marine amphipoda: Gammaridea. British Museum (Natural History), London, United Kigdom.

Lindeque P.K., Parry H.E., Harmer R.A., Somerfield P.J., Atkinson A. (2013). Next generation sequencing reveals the hidden diversity of zooplankton assemblages. PloS One. 8(11), e81327. https://doi.org/10.1371/journal.pone.0081327.

Lobo J., Shokralla S., Costa M.H., Hajibabaei M., Costa F.O. (2017). DNA metabarcoding for high-throughput monitoring of estuarine macrobenthic communities. Scientific Reports. 7, 15618. https://doi.org/10.1038/s41598-017-15823-6.

Mallela J., Milne B.C., Martinez-Escobar D. (2017). A comparison of epibenthic reef communities settling on commonly used experimental substrates: PVC versus ceramic tiles. Journal of Experimental Marine Biology and Ecology. 486, 290-295. https://doi.org/10.1016/j.jembe.2016.10.028.

Marraffini M.L., Ashton G.V., Brown C.W., Chang A.L., Ruiz G.M. (2017). Settlement plates as monitoring devices for non-indigenous species in marine fouling communities. Management of Biological Invasions. 8(4), 559-566. https://doi.org/10.3391/mbi.2017.8.4.11.

McGee K.M., Robinson C.V., Hajibabaei M. (2019). Gaps in DNA-based biomonitoring across the globe. Frontiers in Ecology and Evolution. 7, 337. https://doi.org/10.3389/fevo.2019.00337.

Moura A., Cancela da Fonseca L., Cúrdia J., Carvalho S., Boaventura D., Cerqueira M., Leitão F., Santos M.N., Monteiro C.C. (2008). Is surface orientation a determinant for colonisation patterns of vagile and sessile macrobenthos on artificial reefs? Biofouling. 24(5), 381-391. https://doi.org/10.1080/08927010802256414.

Naylor E. (1972). British Marine Isopods. Keys and Notes for the Identification of the Species– Synopses of the British Fauna. Academic Press, London-New York.

Pacheco A.S., Laudien J., Thiel M., Heilmayer O., Oliva M. (2010). Hard-bottom succession of subtidal epibenthic communities colonizing hidden and exposed surfaces off northern Chile. Sciencia Marina. 74, 147–154. https://doi.org/10.3989/scimar.2010.74n1147.

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

Piñol J., Mir G., Gomez-Polo P., Agusti N. (2015). Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. Molecular Ecology Resources. 15, 819–830. https://doi.org/10.1111/1755-0998.1235.

Porter T.M., Hajibabaei M. (2018). Scaling up: A guide to high-throughput genomic approaches for biodiversity analysis. Molecular Ecology. 27(2), 313-338. https://doi.org/10.1111/mec.14478.

Porter T.M., Hajibabaei M. (2020). Putting COI Metabarcoding in Context: The Utility of Exact Sequence Variants (ESVs) in Biodiversity Analysis. Frontiers in Ecology and Evolution. 8, 248. https://doi.org/10.3389/fevo.2020.00248.

Quast C., Pruesse E., Yilmaz P., Gerken J., Schweer T., Yarza P., Peplies J., Glockner F.O. (2013). The SILVA ribossomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research. 41(D1), D590-D596. https://doi.org/10.1093/nar/gks1219.

Ratnasingham S. (2019). mBRAVE: The Multiplex Barcode Research And Visualization Environment. Biodiversity Information Science and Standards. 3, e37986. https://doi.org/10.3897/biss.3.37986.

Saldanha L. (1995). Fauna submarina Atlântica: Portugal continental, Açores, Madeira. 3ª Edição. Publicações Europa-América. Lisboa.

Schroeder A., Stankovic D., Pallavicini A., Gionechetti F., Pansera M., Camatti E. (2020). DNA metabarcoding and morphological analysis-Assessment of zooplankton biodiversity in transitional waters. Marine Environmental Research. 160, 104946. https://doi.org/10.1016/j.marenvres.2020.104946.

Serrana J.M., Miyake Y., Gamboa M., Watanabe K. (2019). Comparison of DNA metabarcoding and morphological identification for stream macroinvertebrate biodiversity assessment and monitoring. Ecological Indicators. 101, 963-972. https://doi.org/10.1016/j.ecolind.2019.02.008.

Śmietanka B., Burzyński A. (2017). Disruption of doubly uniparental inheritance of mitochondrial DNA associated with hybridization area of European *Mytilus edulis* and *Mytilus trossulus* in Norway. Marine Biology. 164, 209. https://doi.org/10.1007/s00227-017-3235-5.

Śmietanka B., Zbawicka M., Wolowicz M., Wenne R. (2004). Mitochondrial DNA lineages in the European populations of mussels (*Mytilus* spp.). Marine Biology. 146, 79–92. https://doi.org/10.1007/s00227-004-1418-3.

Sokolowski A., Ziólkowska M., Balazy P., Plichta I., Kukliński P., Mudrak-Cegiolka S. (2017). Recruitment pattern of benthic fauna on artificial substrates in brackish low-diversity system (the Baltic Sea). Hydrobiologia. 784(1), 125-141. https://doi.org/10.1007/s10750-016-2862-z.

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. https://doi.org/10.1016/j.seares.2014.01.007.

Steyaert M., Priestley V., Osborne O., Herraiz A., Arnold R., Savolainen V. (2020). Advances in metabarcoding techniques bring us closer to reliable monitoring of the marine benthos. Journal of Applied Ecology. 5(11), 2234-2245. https://doi.org/10.1111/1365-2664.13729.

Taberlet P., Coissac E., Pompanon F., Brochmann C., Willerslev E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. Molecular Ecology. 21, 2045-2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x.

Teixeira M.A.L., Vieira P.E., Pleijel F., Sampieri B.R., Ravara A., Costa F.O., Nygren A. (2019). Molecular and morphometric analyses identify new lineages within a large *Eumida* (Annelida) species complex. Zoologica Scripta. 49, 222-235. https://doi.org/10.1111/zsc.12397.

Thrush S.F., Dayton P.K. (2002). Disturbance to marine benthic habitats by trawling and dredging: implications for marine biodiversity. Annual Review of Ecology and Systematics. 33, 449–473. https://doi.org/10.1146/annurev.ecolsys.33.010802.150515.

Thrush S.F., Pridmore R.D., Hewitt J.E. (1994). Impacts on soft sediment macrofauna: the effects of spatial variation on temporal trends. Detecting Ecological Impacts. 4, 31–41. https://doi.org/10.1016/B978-012627255-0/50006-9.

Underwood A.J., Chapman M.G. (2006). Early development of subtidal macrofaunal assemblages: Relationships to period and timing of colonization. Journal of Experimental Marine Biology and Ecology. 330, 221–233. http://dx.doi.org/10.1016/j.jembe.2005.12.029.

Uribe R.A., Ortiz M., Macaya E.C., Pacheco A.S. (2015). Successional patterns of hard-bottom macrobenthic communities at kelp bed (*Lessonia trabeculata*) and barren ground sublittoral systems. Journal of Experimental Marine Biology and Ecology. 472, 180–188. https://doi.org/10.1016/j.jembe.2015.08.002.

von Ammon U., Wood S.A., Laroche O., Zaiko A., Tait L., Lavery S., Inglis G.J., Pochon X. (2018). Combining morpho-taxonomy and metabarcoding enhances the detection of non-indigenous marine pests in biofouling communities. Scientific Reports. 8(1), 1-11. https://doi.org/10.1038/s41598-018-34541-1.

Wahl M., Link H., Alexandridis N., Thomason J. C., Cifuentes M., Costello M. J., da Gama B. A. P., Hillock K., Hobday A. J., Kaufmann M. J., Keller S., Kraufvelin P., Krüger I., Lauterbach L., Antunes B. L., Molis M., Nakaoka M., Nyström J., bin Radzi Z., Stockhausen B., Thiel M., Vance T., Weseloh A., Whittle M., Wiesmann L., Wunderer L., Yamakita T., Lenz M. (2011). Re-structuring of marine communities exposed to environmental change: a global study on the interactive effects of species and functional richness. PLoS One. 6(5), e19514. https://doi.org/10.1371/journal.pone.0019514.

Warnes G.R., Bolker B., Bonebakker L., Gentleman R., Huber W., Liaw A., Lumley T., Maechler M., Magnusson A., Moeller S., Schwartz M., Venables B. (2020). gplots: Various R programming tools for plotting data. R package version 3.0.3. https://CRAN.R-project.org/package=gplots.

Warwick R.M., Clarke K.R. (1995). New 'biodiversity' measures reveal a decrease in taxonomic distinctness with increasing stress. Marine Ecology Progress Series. 129, 301–305. https://doi.org/10.1111/j.1365-2656.2005.01004.x.

Yu D.W., Ji Y., Emerson B.C., Wang X., Ye C., Yang C., Ding Z. (2012). Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. Methods in Ecology and Evolution. 3, 613–623. https://doi.org/10.1111/j.2041-210X.2012.00198.x.

Chapter 5

ARMS embracing ASMS to complement DNA metabarcoding-based monitoring of macrozoobenthos in NW Iberian coast

Abstract

Autonomous Reef Monitoring Structures (ARMS) have been successfully employed, in combination with DNA metabarcoding, as a standard and replicable framework for marine biodiversity monitoring. However, the impact of the shape of these artificial substrates in the preference of colonizing species has not been addressed yet. Here we use DNA metabarcoding to compare the overtime macroobenthic species colonization between ARMS and Artificial Seaweeds Monitoring System (ASMS), a substrate that mimics the structure of a seaweed, thereby very distinct from the 9 superimposed PVC plates that compose ARMS. To this end, we deployed both substrates in two different locations (Ría de Vigo and Ría de Ferrol, NW Iberian coast) and collected them after 6, 9 and 12 months to assess the species composition of the the colonizing communities. A total of 218 macrozoobenthic species were detected. We recorded a large fraction exclusive for a particular substrate (43.1% of the total taxa detected), and no species was detected consistently in all substrates/sampling times/sites combinations. Although we observed a higher number of taxa in ASMS at both sampling sites, the two substrates were quite complementary in their macrozoobenthic species composition, with only approximately 30% of species were overall recorded in both substrates. The majority of the species were detected exclusively in one substrate type, and some taxonomic groups as a whole appear to occur preferentially in ASMS (e.g. Crustacea and Gastropoda) or ARMS (e.g. Annelida, Echinodermata and Porifera). Hence, the shape of the substrate strongly affected the colonization preferences, indicating that important fractions of the regional diversity may be overlooked if only one substrate type is used for monitoring in this region. Overall, compared to Ría de Ferrol, in Ría de Vigo we consistently recorded a moderately higher number of species most of them were exclusive to this location. Thus, we demonstrate that despite the custumary use of ARMS for macrozoobenthos monitoring, using ASMS we complemented the recovery of species and enlarged the scope of the phylogenetic diversity recorded. In future monitoring studies in the region, the simultaneous use of both substrates as sampling devices is expected to provide a much more comprehensive profile of the macrozoobenthos and to minimize overlooked taxa.

Keywords: Artificial Reef Monitoring Structures (ARMS); Artificial Seaweed Monitoring System (ASMS); Coastal biodiversity; Coastal monitoring; NW Iberian coast.

5.1 Introduction

Species interactions within marine communities are responsible for the maintenance of a biological network (i.e. consumers, predators and decomposers) highly important in ecosystem processes (e.g. energy flow, primary production, nutrient recycling; Doney et al., 2011; Niiranen et al., 2013). However, the functioning and ability of marine ecosystems to provide services can be severely compromised due to the effects of global impacts (e.g. multiple stressors, other human pressures; Burrows et al., 2011; Duarte, 2014), which can have a high impact in enclosed and semi-enclosed basins (Danovaro, 2003).

The Lusitanian biogeographic province (Spalding et al., 2007) constitutes a unique spot for marine research, since it harbors a high diversity of macrofauna from various adjacent regions, and many species have their northern or southern range limits in this area (Pereira et al., 2006). Monitoring these communities is particularly relevant to assess the impact of global change on marine biodiversity and ecosystems (e.g. shifts on species ranges expansions or alteration of dispersal patterns) and changes on species interactions (Doney et al., 2011; Henriques et al., 2014).

However, large-scale biodiversity assessment and hard-bottom communities sampling is challenging, mostly due to difficulties in access and diving limitations (e.g. scuba diving costs; Bianchi et al., 2004; Borja et al., 2016). Implementing innovative and standardized methods is essential (Danovaro et al., 2016) to retain and make data accessible for analysis in a wide use and re-use application in further studies (Tanhua et al., 2019). Technical advances in monitoring approaches, through the implementation of innovative molecular approaches (Steyaert et al., 2020), namely DNA metabarcoding, provide an opportunity to rapidly improve the accuracy and throughput for marine biodiversity assessment and monitoring (Bourlat et al., 2013; Cahill et al., 2018). Furthermore, assessing marine ecosystems in an integrative way, through the use of artificial substrates coupled with DNA metabarcoding may be a valuable alternative as a replicable and standard methodology of marine macrozoobenthic monitoring, as simple and cost-effective as currently possible. Although artificial substrates have already been used to promote colonization and monitor marine communities (e.g. Pearman et al., 2019; Ros et al., 2013; Sedano et al., 2020), the implementation of such strategy in large-scale comparisons is difficult due to the low level of standardization of the methodologies (i.e. materials with different size and composition, and different sampling times and processing protocols).

Autonomous Reef Monitoring Structures (ARMS) originally developed to mimic coral reefs diversity (Zimmerman and Martin, 2004), have a structure with cavities influenced by high and low light spaces and

various flow regimes (Leray and Knowlton, 2015). These characteristics provide shelter for small invertebrates (e.g. protecting against predation) and surfaces for sessile organism's settlement (Leray and Knowlton, 2015; Ransome et al., 2017). Deployed over long-term, ARMS allow to assess and interpret the diversity, distribution and structure of hard-bottom marine communities (Templado et al., 2010), and has been frequently applied in the assessment of diversity in a variety of geographic regions (Caribbean and Indo-Pacific – Plaisance et al., 2011; Singapore – Chang et al., 2020; French Polynesia – Ransome et al., 2017; Red Sea – Al-Rshaidat et al., 2016, Carvalho et al., 2019; Adriatic Sea – Pennesi and Danovaro, 2017; Atlantic coast – Leray and Knowlton, 2017; Europe - Obst et al., 2020; Iberian Coast – David et al., 2019). Artificial Seaweeds Monitoring System (ASMS) is an alternative artificial substrate, which was employed in a study developed in Ría de Ferrol (NW Iberian Peninsula; Carreira-Flores et al., 2020) to mimic macroalgae. Since ASMS have a different tridimensional structure from ARMS, can attract different species, especially highly-mobile fauna.

Systematic sampling, using ARMS and ASMS as monitoring tools with the available and cost-efficient high-throughput sequencing (i.e. DNA metabarcoding) will allow for an accurate and comparable assessment of a wide spectrum of the biodiversity of hard-bottom communities. Despite the increasing implementation of ARMS in different geographical locations for hard-bottom marine monitoring using molecular approaches (Chang et al., 2020; Obst et al., 2020; Pearman et al., 2020), the comparison between both artificial substrates and their influence on the assessment of macroinvertebrate species has not been performed yet. Given the results obtained in chapter 4, showing the aptitude of artificial substrates combined with DNA metabarcoding for efficient benthic monitoring, we designed a long-term monitoring strategy employing ARMS and ASMS to investigate the influence of tridimensional structures on macrozoobenthic colonization, examing whether both substrates support similar assemblages or if they supply complementary information about species occurrence. For this purpose, we selected two sites in NW Iberian coast to deploy the artificial substrates, and assessed overtime patterns of colonization.

5.2 Materials and methods

5.2.1 Study area

This study was carried out along NW Atlantic Iberian coast at two different locations (Fig. 5.1): Bajo Tofiño (42°13'42.3"N 8°46'43.2"W, Ría de Vigo, Spain) and San Cristovo (43°27'53.8"N 8°18'00.7"W, Ría de Ferrol, Spain).

Ría de Ferrol is a fully marine environment structurally composed by a semi-enclosed bay connected to the Atlantic Ocean by a narrow channel (Moreira et al., 2009; Urgorri et al., 1992). As in Ría de Vigo (see 3.3.1 Materials and Methods – sampling design section for more details), Ría de Ferrol is composed by busy ports and directly subjected to human perturbations (e.g. sewage runoff or harvesting; Barroso et al., 2000; Prego et al., 2008). The two locations present a semidiurnal tidal regime, and the wave regime is dominated by winds from NW and during autumn-winter seasons occurs most of the storms (Rubal et al., 2011; Dias et al., 2002).

We also deployed substrates at Viana do Castelo (41°40'45.2"N 8°51'44.4"W, Portugal), however due to atmospheric conditions and the natural characteristics of the sampling zone, and despite the effort to find the substrates, they were buried in the sand and not used in further analysis.

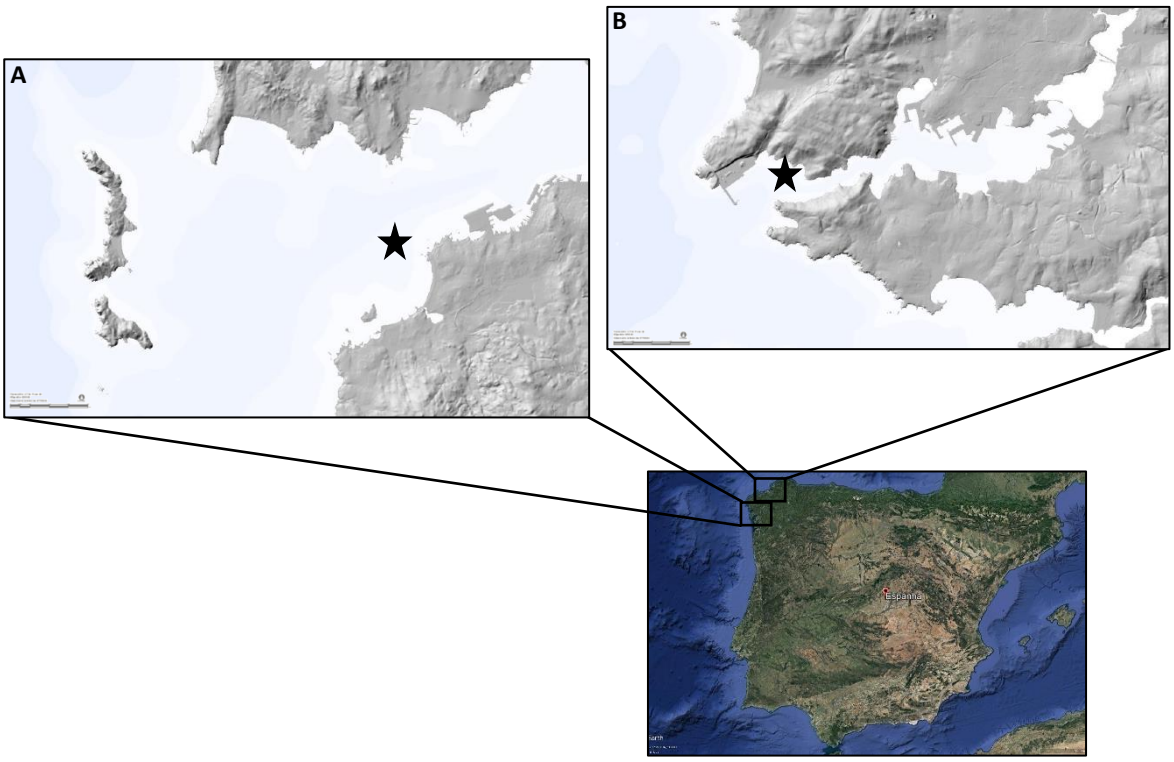


Figure 5.1. Study sites (black star): **A** – Bajo Tofiño, Ría de Vigo, Spain; **B** – San Cristovo, Ría de Ferrol, Spain.

5.2.2 Sampling design

We selected two types of artificial substrates: a) ARMS, which are small tiered platforms, composed by nine piled up plates (23 x 23 cm) of grey Type I PVC separated by spacers affixed to the seafloor; b) ASMS, which are plastic commercial artificial plants (IKEA, Sweden), with 28 cm height and composed by green polyethylene with a complex structure formed by different orientation of the plant branches.

Three replicates of the two selected substrates (Fig. 5.2) were deployed in June 2018 anchored to a cement plate (60 x 60 cm) and affixed to the bottom (approximately 11 m of depth). After 6, 9 and 12 months of deployment the substrates were removed from Ría de Vigo and Ría de Ferrol.

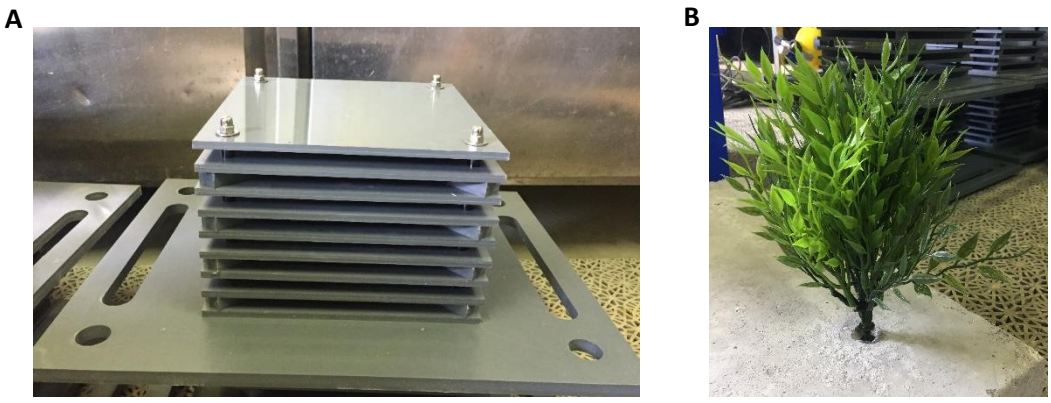


Figure 5.2. Artificial substrates used for marine macrozoobenthic colonization. A – ARMS; B – ASMS.

5.2.3 ARMS and ASMS recovery and processing

In order to limit the loss of motile organisms, divers enclosed the ARMS in a labeled plastic box and lifted to the boat (Fig. 5.3). For ASMS, each sample were carefully enclosed in a 500 µm mesh bag and then introduced in a hermetic plastic bag before being released from the substratum with a scraper (Fig. 5.4). This is important to prevent the escape of small motile organisms associated to the substrate.

At the laboratory, samples were photographed and then processed. We disassembled ARMS plate by plate following Leray and Knowlton (2015), and each branch of ASMS was separated. Then, samples were carefully washed using filtered seawater and the mobile and sessile fauna were separated. While the mobile fauna was brushed and sieved (500 µm), the sessile fauna was scraped with a spatula into a tray. All

samples were then preserved in ethanol and stored at -20 ^{0}C until further analysis. The water in the container of each substrate was sieved, and the retained organisms were preserved with mobile fauna.

Figure 5.3. ARMS recovery procedure.

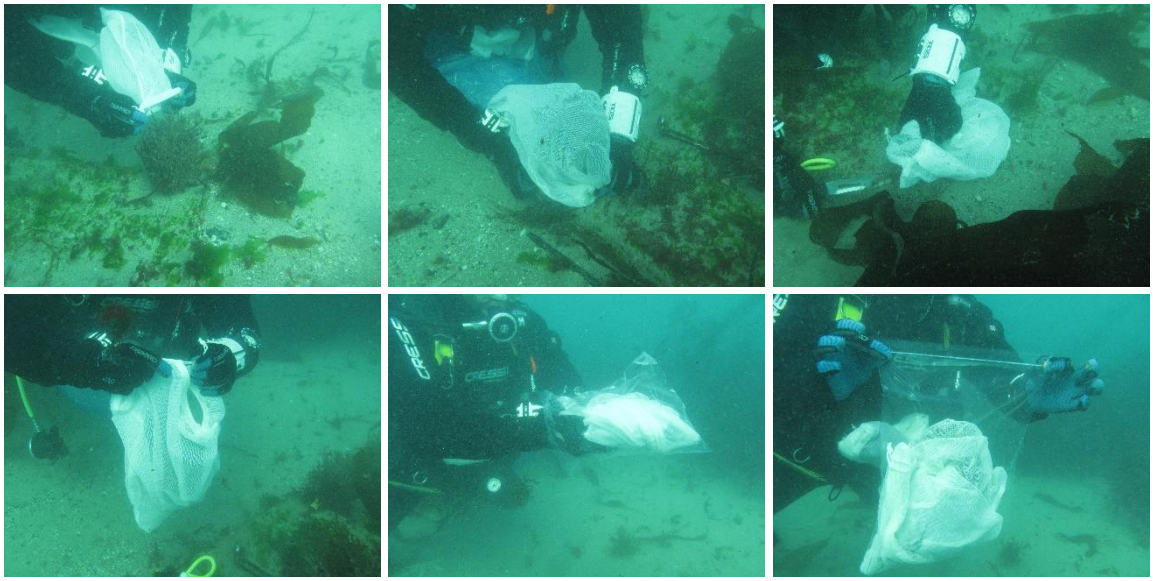


Figure 5.4. ASMS recovery procedure.

5.2.4 DNA metabarcoding

DNA extraction procedures were adapted from Ivanova et al. (2006). Based on the wet weight of each sample (Braukmann et al., 2019), we added an appropriate volume of lysis buffer solution (100 mM NaCL, 50mM Tris-HCL pH 8.0, 10 mM EDTA, 0.5% SDS) and we incubated the samples at 56 ºC overnight and 200 rpm. To maximize diversity recovery, two-aliquots of each lysate was used, totalling two DNA extractions per sample. After extraction the aliquots of genomic DNA for the same sample were pooled in a single microtube and sent for high-througput sequencing (HTS).

The production of amplicon libraries and the HTS were carried out at Genoinseq (Cantanhede, Portugal), as described below. One primer-pair targeting the COI-5P barcode region and one primer-pair targeting the 18S rRNA gene were selected to amplify the marine macrozoobenthic communities from each sample (Table 5.1). Each PCR reaction were performed for each sample using KAPA HIFI HotStart PCR Kit according to manufacturer instructions, 0.3 µM of each PCR primer and 50 ng of template DNA in a total volume of 25 µL (PCR conditions are displayed in Table 5.1). Negative controls for DNA extractions and PCR, and positive PCR controls were included in the molecular analysis workflow. We also tested the amplification of genomic DNA with another primer targeting COI-5P (LCO1490/Ill_C_R; Shokralla et al., 2015), however with no success.

Second PCR reactions added indexes and sequencing adapters to both ends of the amplified target region according to manufacturer's recommendations (Illumina, 2013). PCR products were then one-step purified and normalized, pooled and pair-end sequenced in an Illumina MiSeq® sequencer with the V3 chemistry, according to manufacturer's instructions (Illumina, San Diego, CA, USA).

The ARMS samples from Ría de Ferrol after 6 months of deployment, and from Ría de Vigo of sessile fauna after 6 and 9 months of deployment, and the ASMS samples from Ría de Vigo of mobile fauna after 6 months of deployment did not amplify for COI-5P barcode region and were not included in further analyses.

Table 5.1. Primer-pairs and respective thermal cycling conditions used to amplify marine macrozoobenthic communities. F – Forward; R – Reverse; bp – Base pairs.

Primer-pair	Direction (5'-3')	Reference	PCR thermal cycling conditions
mlCOIintF LoboR1 (313 bp)	(F) GGWACWGGWTGAACWGTWTAYCCYCC (R) TAAACYTCWGGRTGWCCRAARAAYCA	Leray et al., 2013 Lobo et al., 2013	(1) 95 °C (3 min); (2) 35 cycles: 98 °C (20 s), 60 °C (30 s), 72 °C (30 s); (3) 72 °C (5 min).
TAReuk454FWD1 TAReukREV3 (400 bp)	(F) CCAGCASCYGCGGTAATTCC (R) ACTTTCGTTCTTGATYRA	Stoeck et al., 2010; Lejzerowicz et al., 2015	(1) 95 °C (3 min); (2) 10 cycles: 98 °C (20 s), 57 °C (30 s), 72 °C (30 s); (3) 25 cycles: 98 °C (20 s), 47 °C (30 s), 72 °C (30 s); (4) 72 °C (5 min).

5.2.5 Data processing

DNA sequences were quality and size filtered to remove sequencing adapters (PRINSEQ v.0.20.4, Schmieder and Edwards, 2011). The forward-R1 and reverse-R2 reads generated were merged by overlapping pair-end reads using mothur v.1.39.5 (Schloss et al., 2009; Kozich et al., 2013). Then, we remove primers and determine a minimum sequence length of 150 base pairs. The usable reads were then submitted to mBrave (www.mbrave.net, Ratnasingham, 2019) and SILVAngs database (https://ngs.arb-silva.de/silvangs/, Quast et al., 2013), to generate the OTU tables and taxonomic assignments for COI and 18S, respectively. In both cases, taxonomic assignments were attributed when displaying ≥97% similarity with reference sequences. Only reads with match at species-level were used for further analysis, and singletons and rare OTUs (i.e. less than 8 sequences) were discarded (Lobo et al., 2017). Any read matching to non-metazoan was also excluded. The validity of the species names was then verified in the World Register of Marine Species (WoRMS) database (http://www.marinespecies.org/).

5.2.6 Statistical analyses

The proportion of overlapping and unique species detected between substrates was determined for both sampling sites and displayed using Venn diagrams, obtained with the R package VennDiagram (Chen and Boutros, 2011); while qualitative data of species distribution among taxonomic groups was displayed through bar graphs (GraphPad Software, Inc.).

Multivariate analyses were carried out considering presence/absence of the taxa due to the qualitative nature of the molecular data. We performed all the community analysis in PRIMER v6.1.1.16 (Primer-E Ltd, Plymouth, UK), except for the mentioned situations. Bray-Curtis measure of similarity for presence/absence of species was used to compare the two fractions of fauna (mobile and sessile), and to investigate differences between the substrates, study sites and sampling times. Data were visually explored using hierarchical clustering (CLUSTER; linkage method: UPGMA) to investigate groups of samples on macrozoobenthic community structure for each substrate among sampling times and between study sites, and for the two fractions of fauna. One-way analyses of similarities (ANOSIM) tests were used to look for differences in the colonized community between substrates and among the distinct fauna fractions. Similarity Percentages Test (SIMPER) was used to test the hypothesis about differences in the composition of the macrozoobenthic communities in the substrates among all sampling times and in each study site, and to identify the taxa which most contributed for group differences. To estimate taxonomic diversity, a hierarchical taxonomic classification was used to represent the relatedness of individual species based on the composition detected. From the five taxonomic levels compiled (species, genus, family, order, class and phylum), the Average Taxonomic Distinctiveness (AvTD - Δ^+; Clarke and Warwick, 1998), were calculated based on the average pair-wise path lengths for each sample (i.e. substrate/sampling time/site) to evaluate the assemblages diversity, taking into account Linnean taxonomic distance among the species from a sample.

Non-metric multidimensional scaling (nMDS) analyses were performed based on Bray-Curtis resemblance coefficient between samples to visualize community distribution from the distinct fauna fractions (mobile and sessile) and from the two sampling sites (Ría de Vigo and Ría de Ferrol) for all substrates among sampling times (PAST v4.03; Hammer et al., 2001).

Species occurrence was assessed for the consideration of their exclusivity (i.e. species only detected in one substrate/sampling time/site combination), partial exclusivity (i.e. species present in one sampling time and in both sampling sites, independently of the substrates), partial pervasiveness (i.e. species

detected among all sampling times, but only in one sampling site, independently of the substrate), and pervasiveness (i.e. species shared among both substrates and all sampling times and sites).

Heat map was built with the R package "gplots" (Warnes et al., 2020) using the presence/absence data of the species which mainly account for the SIMPER differences detected, in the different substrates and sampling times.

5.3 Results

A total of 218 species, representing 12 different phyla (Annelida, Arthropoda, Bryozoa, Chordata, Cnidaria, Echinodermata, Mollusca, Nematoda, Nemertea, Platyhelminthes, Porifera and Sipuncula; Table S.5.1) were identified in all combinations of sampling times and sites in both ARMS and ASMS. Through the observation of photographs, differences between sampling sites for both substrates are patent (Figs. 5.5 and 5.6), Additionally, slight differences between sampling times within each study site were observed, except in 12 months ASMS from Ría de Vigo (Fig. 5.6).

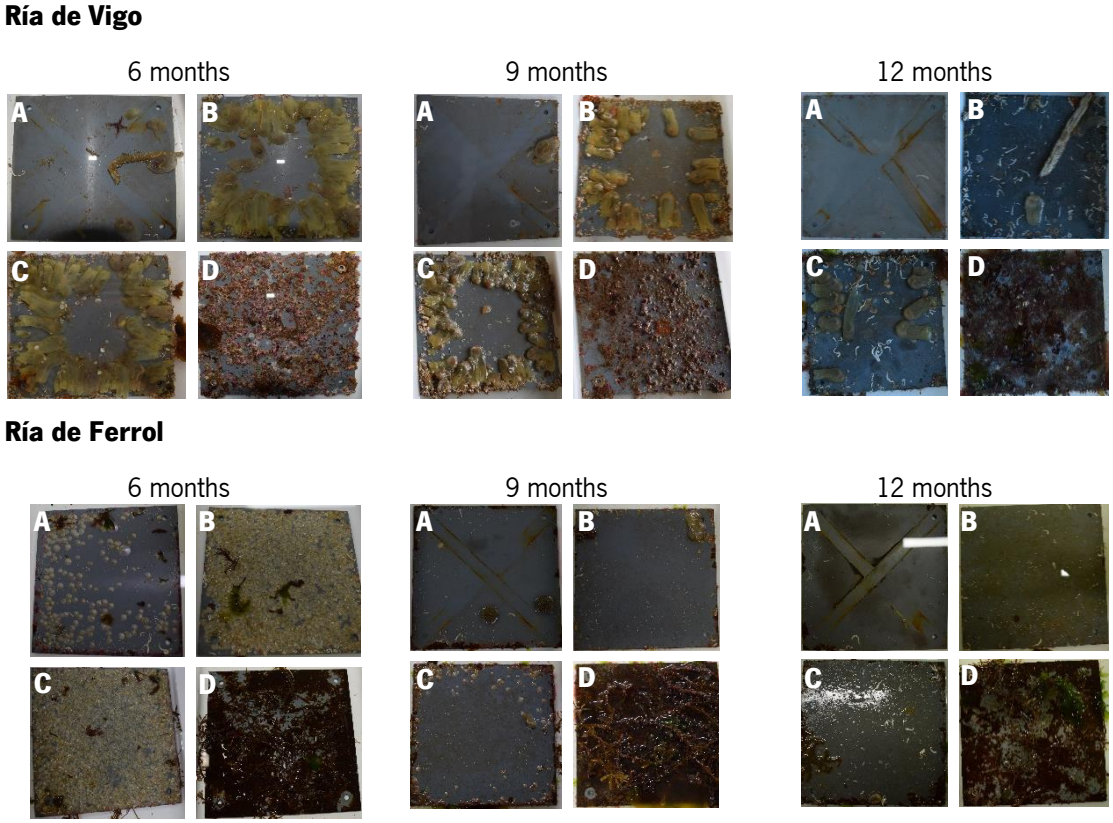


Figure 5.5. Sampled ARMS face plates collected after 6, 9 and 12 months of deployment at Bajo Tofiño - Ría de Vigo, and San Cristovo - Ría de Ferrol. A – Plate 1 top; B – Plate 5 bottom; C – Plate 9 top; D – Plate 9 bottom.

Ría de Vigo

6 months

9 months

12 months

Ría de Ferrol

6 months

9 months

12 months

Figure 5.6. ASMS collected after 6, 9 and 12 months of deployment at Bajo Tofiño - Ría de Vigo and San Cristovo - Ría de Ferrol.

High-throughput sequencing from marine macrozoobenthic samples, for both markers and for the total of 24 samples, generated a total of 1,348,329 usable reads (Table S.5.2) of these, 49% were assigned to marine macrozoobenthos species (30% using mlCOIintF/LoboR1 and 19% with TAReuk454FWD1/TAReukREV3). Of the remaining reads, 2% were singletons or rare sequences (<8 reads) and 49% could not be assigned to macrozoobenthic species.

Based on the sequences identified at species-level, 14 high-rank taxa (i.e. phylum and sub-phylum level) were retrieved in ARMS samples, where Annelida (31 species), Crustacea (21 species), Echinodermata and Hydrozoa (19 species for both) were the most well represented. The major contributors for ASMS community diversity were Crustacea (36 species), Hydrozoa (29 species) and Annelida (19 species) (Fig. 5.7). Furthermore, four more taxa (Isopoda, Pycnogonida, Nematoda and Sipuncula) were exclusively detect in ASMS.

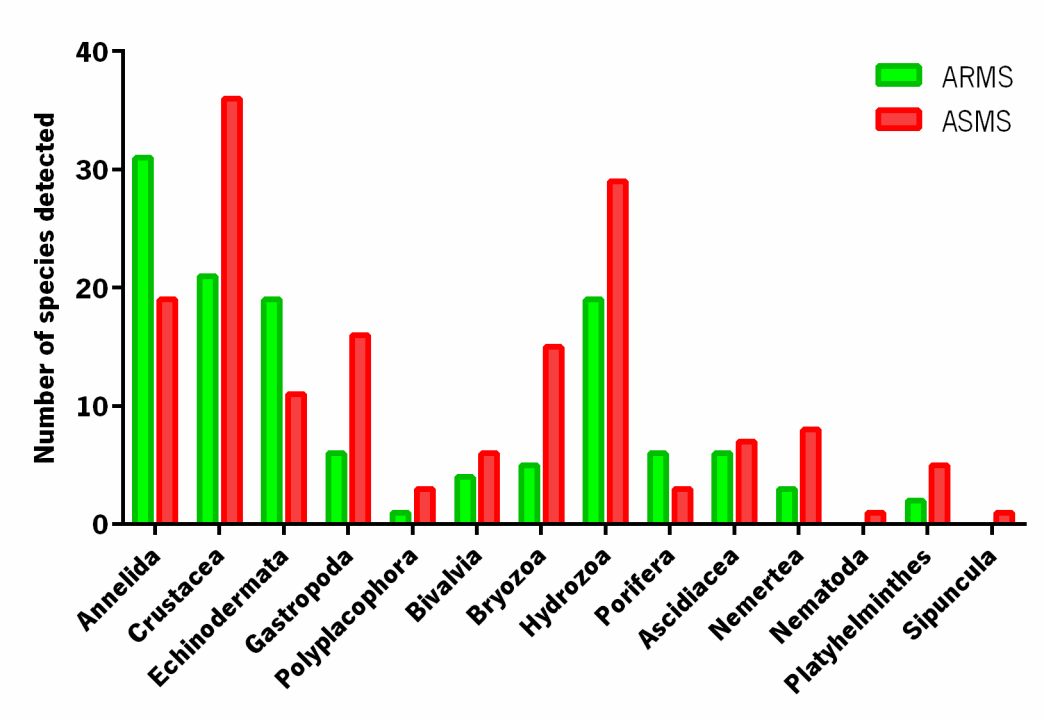


Figure 5.7. Taxonomic distribution of the total species detected in each substrate: ARMS and ASMS.

For ARMS, the highest number of species detected was recorded in Ría de Vigo, after 9 months of deployment (49 species). Although the lowest detection of species was verified in Ría de Ferrol, after 6 months of deployment (18 species), we detected more taxonomic groups at this sampling time than in others. For ASMS, the highest and lowest number of species were detected in Ría de Vigo, after 12 months (80 species) and 6 months (44 species) of deployment, respectively (Fig. 5.8). Lower taxonomic diversity was also detected for the same ASMS with the lowest number of species retrieved.

For the total species detected on each substrate, the similarity within substrates was low, especially for ARMS (27.55%). Additionally, high dissimilarity was displayed between both substrates, regardless sampling time and sites (Table 5.2). The community composition and taxonomic groups detected were different between substrates, and significant differences were found between substrates across sampling sites (R = 0.574, significance level = 1%; ANOSIM).

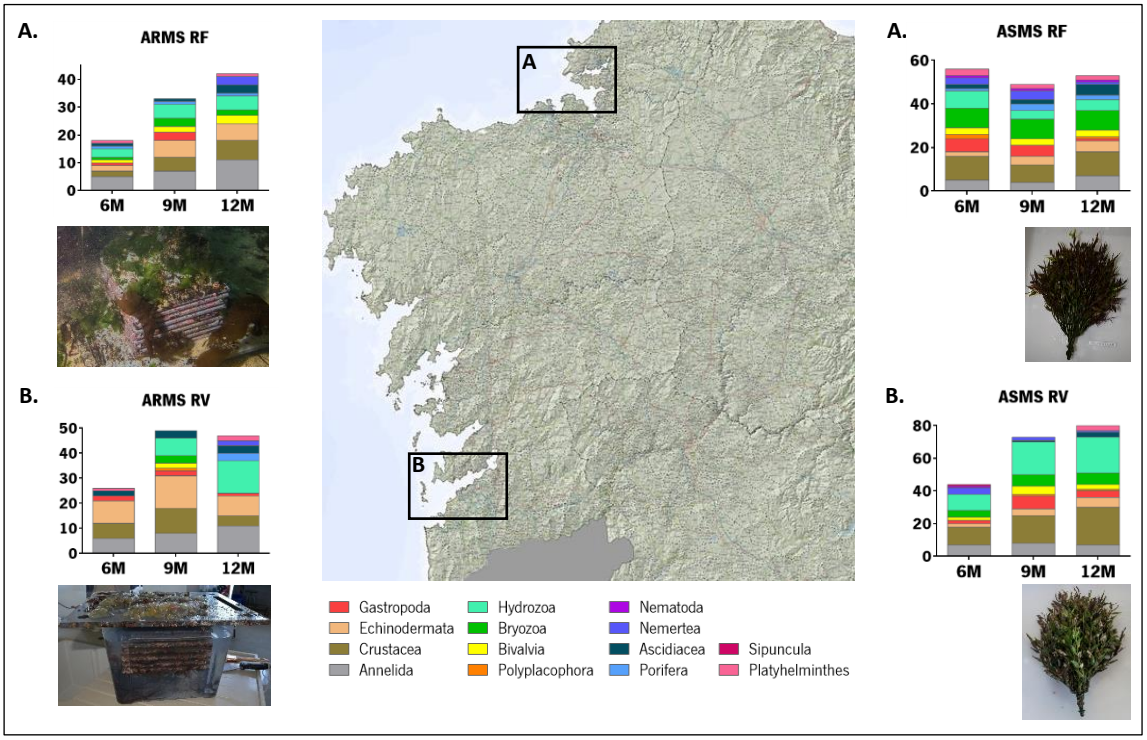


Figure 5.8. ARMS and ASMS substrates from two sampling sites (A) Ría de Ferrol (RF) and (B) Ría de Vigo (RV). Bar charts represents the abundance and taxonomic distribution of species detected on each substrate among sampling times (6M – 6 months; 9M – 9 months; 12M – 12 months).

Table 5.2. SIMPER analysis showing species-level differences in community composition across substrates.

Substrate	Average similarity within substrates	Substrate comparison	Average dissimilarity between substrates
ARMS	27.55	ARMS/ASMS	74.56
ASMS	40.71		

Although ASMS retrieved more taxa in both sampling sites, both substrates were highly complementary in their ability to be colonized by marine macrozoobenthic species. A significant percentage of species (43.6%) were only detected in ASMS, and only 29.8% of the species were detected on both substrates (Fig. 5.9 A and B).

For the samples from Ría de Vigo (Fig. 5.9 A), the majority of the species only detected by one substrate (42 in ARMS and 64 in ASMS) were exclusive species (i.e. species detected only in one substrate/sampling time combination; 79% in ARMS and 56% in ASMS). Furthermore, only 5% of the species in ARMS and 10% of the species in ASMS were ubiquitous. For the common species detected by both substrates in Ría de Vigo, only one species was detected in all substrates/sampling time combinations. The same pattern is even more evident in the samples from Ría de Ferrol (Fig. 5.9 B): 92% of the species only detected in ARMS and 71% in ASMS were exclusive species. While in ASMS 13% of the species were pervasive, in ARMS no species were detected among all sampling times.

By combining the detected species by both substrates on each sampling site, a similar number of species were detected (69.7% RV vs 60.6% RF; Fig. 5.9 C). However, differences in taxonomic groups were recorded. For example, Sipuncula was only detected in samples from Ría de Vigo. For the species only detected in Ría de Vigo, the majority are exclusive species (i.e. species detected only in one substrate/sampling time combination; 57% of the species). The same pattern is observed in the Ría de Ferrol, where 68% were exclusive species, and no species were shared among both substrates and sampling times.

For the common species detected in both sampling sites, 41 were detected in both substrates, while the remaining 25 species were only detected in one substrate. Additionally, 53 of the common species were detected in the same sampling time, of which 23 were detected in both substrates. The remaining 14 species had a variable occurrence among substrates, sampling times and sites (e.g. *Pilumnus hirtellus,* was recovered in Ría de Vigo after 6 months (ARMS) and 9 months (ARMS and ASMS), and after 12 months of deployment in Ría de Ferrol (ARMS).

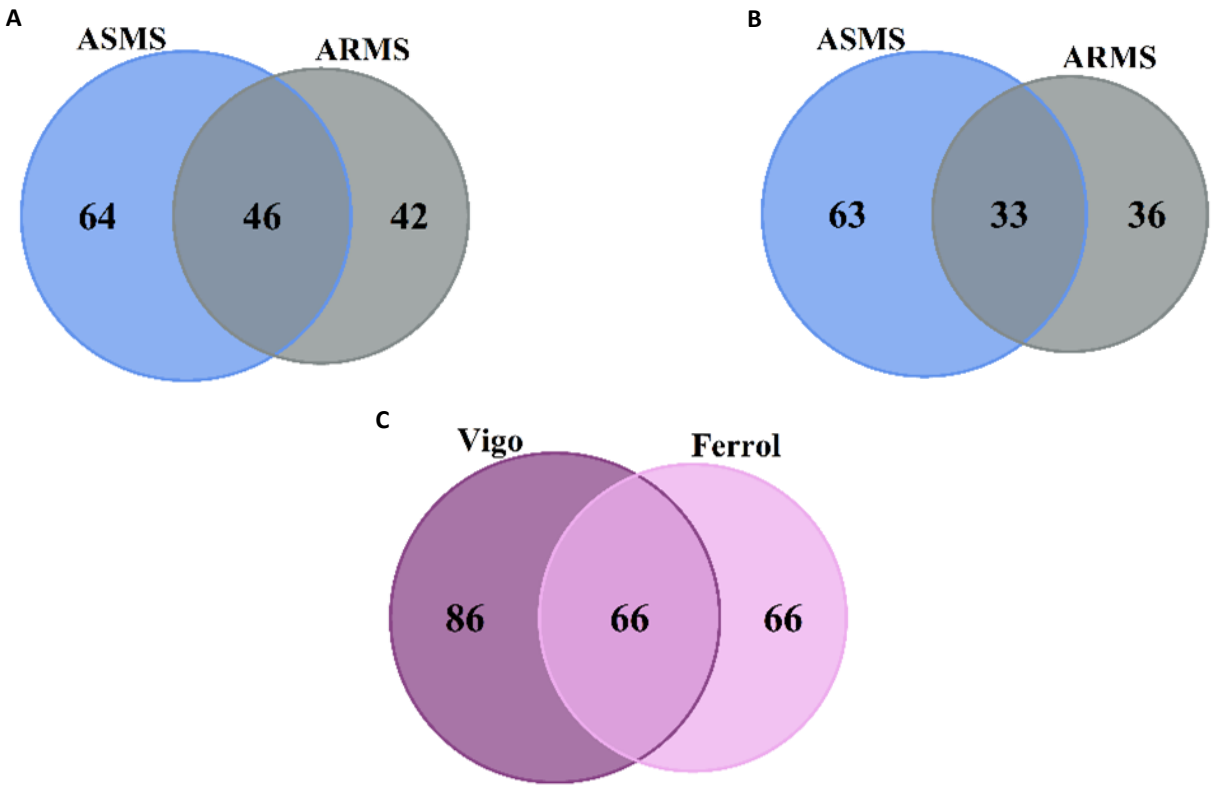


Figure 5.9. Partitioning of the marine macrozoobenthic species detected exclusively by ARMS (grey), exclusively by ASMS (blue), and shared by both substrates (overlapping circles) for all sampling times in (A) Ría de Vigo and (B) Ría de Ferrol; and (C) the combined proportion of species detected by both substrates in each sampling site.

Overall, no species was recorded as completely ubiquitous (detected in all sampling times, substrates and study sites). The prevalence of exclusive species (i.e. species detected only in one substrate/sampling time/site combination) were observed for 43.1% of the total species (94 species), with differences detected in taxonomic groups composition between substrates. Compared to ASMS, Annelida, Echinodermata and Porifera were the taxonomic groups with higher differences among sampling sites and times in ARMS. Whereas Crustacea and Mollusca were the groups that accounted mostly for the differences found in ASMS. Combining both substrates, the taxonomic groups which displayed higher differences in exclusive species between study sites were Bryozoa, Crustacea and Hydrozoa.

Considering all detected species, partial pervasiveness was greater than partial exclusivity, with 17 and 4 species falling under these conditions, respectively. An important fraction of species (47%) had a variable

occurrence among substrates, sampling times and sites (Fig. 5.10), and interesting time-geographical patterns of colonization were observed for some of these species. For example, the caprelid *Caprella acanthifera* detected in Ría de Ferrol after 6 months of deployment, was then detected in Ría de Vigo at 9 and 12 months of deployment. Other example was recorded for three decapods (*Eualus cranchii, E. occultus* and *Hyppolyte varians*) that were first detected after 9 months of deployment in Ría de Vigo, and then after 12 months in Ría de Ferrol.

A SIMPER analysis, comparing the species detected on each substrate across sampling times, revealed low levels of similarity within and high dissimilarity among sampling times. Additionally, substrates displayed high dissimilarity among sampling times, regardless sampling sites (Table 5.3).

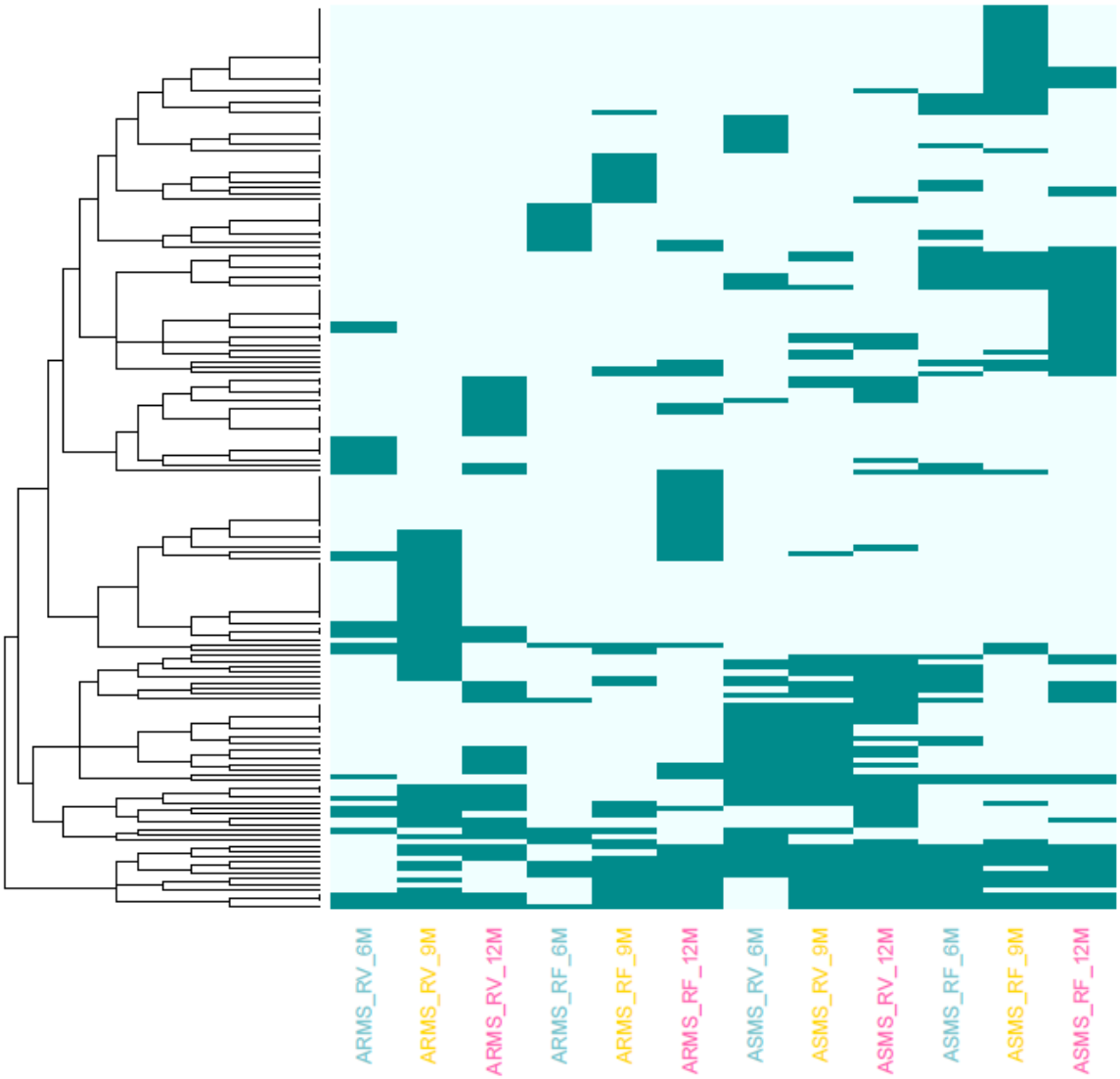


Figure 5.10. Heat-map of marine macrozoobenthic species most accounting for dissimilarities between substrates, after SIMPER analysis. RV - Ría de Vigo; RF - Ría de Ferrol. 6M – 6 months (blue); 9M – 9 months (yellow); 12M – 12 months (pink). Blue represents species detection.

Table 5.3. SIMPER analysis showing species-level differences in community composition among sampling times and between substrates. 6M – 6 months; 9M – 9months; 12M – 12 months.

Substrate	Sampling time	Average similarity within sampling time	Sampling times comparison	Average dissimilarity between sampling times
ARMS	6M	13.64	6M / 9M	71.97
	9M	36.14	6M / 12M	77.81
	12M	28.89	9M / 12M	66.58
ASMS	6M	34.00	6M / 9M	59.11
	9M	29.03	6M / 12M	59.19
	12M	36.36	9M / 12M	53.87
ARMS vs ASMS			6M	83.95
			9M	69.74
			12M	67.03

From the total species recovered, the mobile and sessile fauna had 73 and 82 species detected in ARMS and 90 and 124 species detected in ASMS, respectively. The mobile fraction in ARMS was dominated by annelids (16 species) and echinoderms (17 species), while echinoderms (10 species), gastropods (11 species) and hydrozoans (11 species) were the most representative groups in ASMS. For sessile fauna, in both substrates, annelids and hydrozoans were the most abundant groups. However, mobile fauna samples yielded sessile taxa and vice versa, such as the higher number of annelids detected in sessile fraction than in mobile for both substrates (21 for ARMS and 17 for ASMS).

Cluster (Fig. S.5.1) and non-metric multidimensional analysis (Fig. 5.11), comparing species detection by the two fractions, between substrates and in both sampling sites, showed two distinct clustered groups. Group A combined all samples from sessile fauna with two mobile samples from Vigo (9 and 12 months), and could be divided in A1 - samples from sessile fauna from Ferrol, and A2 - samples from sessile fauna from Vigo, with the two samples from mobile fauna (also from Vigo). We could divide A2 in two subgroups: samples from mobile (A2b) and sessile fauna (A2a) (Fig. S.5.1). The species which most accounted for group A2b are typical macrozoobenthic mobile fauna: one polychaete (*Alentia gelatinosa*), five crustaceans (*Caprella acanthifera, Achelia echinata, Eualus cranchii, Hyppolyte varians* and *Astacilla damnoniensis*)

and one echinoderm (*Ophiocomina nigra*). The other group – B - comprised the remaining samples from mobile fauna, with an average dissimilarity from group A of 77.05%.

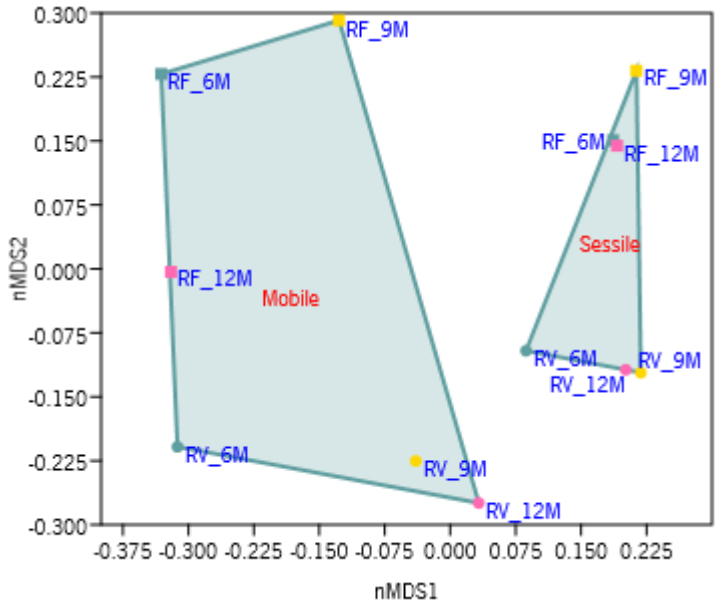


Figure 5.11. Non-metric multidimensional scaling (nMDS), based on Bray-Curtis similarity index, of the species detected in both artificial substrates for mobile and sessile fauna fractions separately. Samples combine all species detected by both substrates on each sampling site (RV – Ría de Vigo; RF – Ría de Ferrol) among sampling times (6M – 6 months, 9M – 9 months and 12M – 12 months).

Seasonal changes on species composition differed between substrates and within sampling sites, where only three taxa were represented among all sampling times and substrates (Annelida, Decapoda, and Echinodermata). For example, in Ría de Vigo Bivalvia was detected in ASMS among all sampling times, while in ARMS was only detected at 9 months of deployment. This pattern differs from samples from Ría de Ferrol, where Bivalvia was detected in both substrates among all sampling times. From the most abundant taxa, taxonomic groups showed fluctuations in species detected and demonstrated differences between substrates colonization (Fig. 5.12). For example, only three taxonomic groups in two substrate/sampling site combinations had highest species detected on winter (6 months of deployment): Amphipoda (ARMS RV and ASMS RF), Gastropoda and Hydrozoa (both from ASMS RF). Only on ARMS from Ría de Vigo, gastropods revealed the highest species detection rate in spring (9 months of deployment).

Combining the total species recovered in Ría de Vigo in both substrates, after 6 months of deployment we detected the lowest number of species (67 species), while after 9 months the highest number of species was recovered (103 species). For Ría de Ferrol, the same number of species was detected after 6 and 9 months (66 species), increasing to the highest number of species recovered after 12 months of deployment (76 species). Comparing the total species detected on each location, the detected taxonomic groups varied across sampling sites. In Ría de Vigo, the majority of taxa had most species detected on spring (after 9 months of deployment), except for Amphipoda (winter – 6 months) and Hydrozoa (summer – 12 months). Similarly, in Ría de Ferrol we also detected more amphipods on winter (after 6 months of deployment).

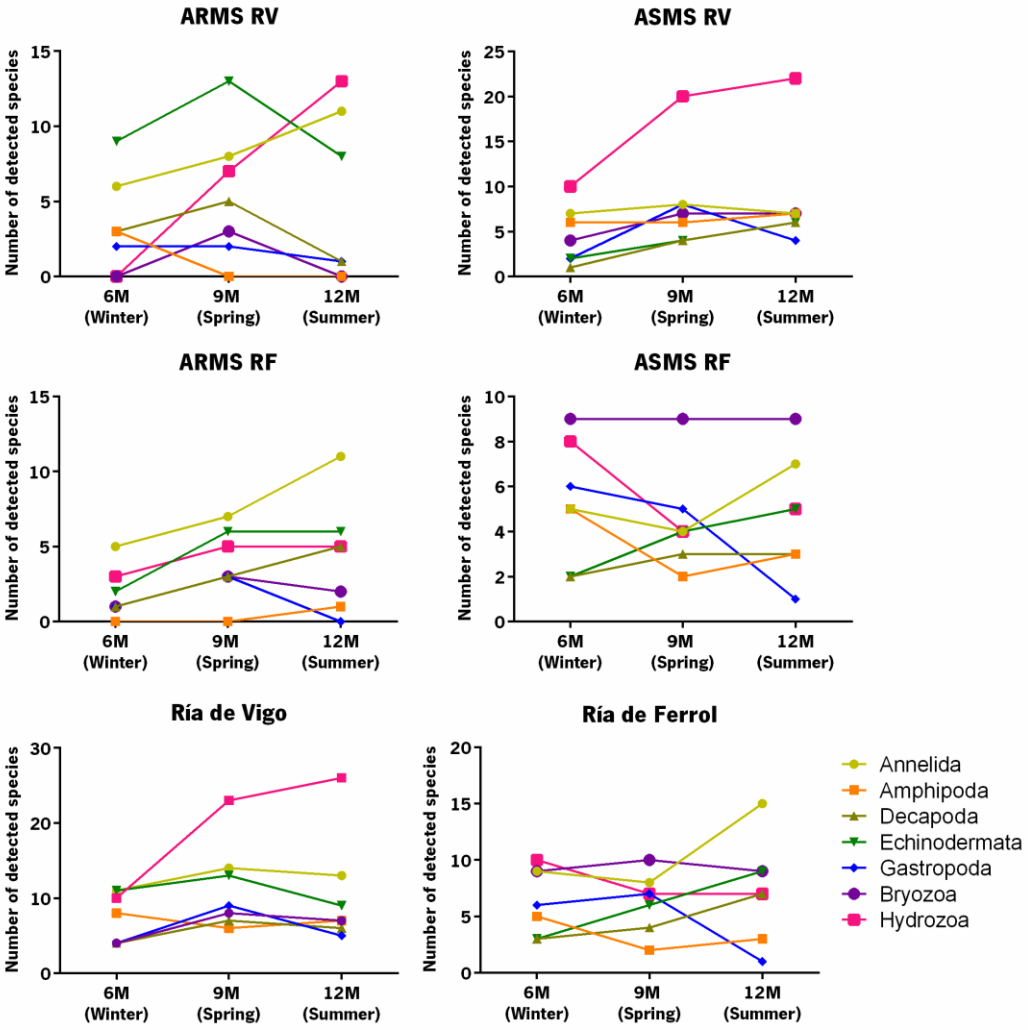


Figure 5.12. Fluctuations in presence/absence of taxa for most abundant taxonomic groups within ARMS and ASMS from Ría de Vigo (RV) and Ría de Ferrol (RF), and for all taxa detected in each Ría divided by sampling seasons (6 months – winter, 9 months – spring, 12 months – summer).

The AvTD funnel plot (AvTD = 95.5 ± 0.8; Fig. 5.13 and Table S.5.3) showed that most of the samples match to the expected distribution in the 95% confidence interval, suggesting a good degree of taxonomic stability. Only two samples, ASMS from Ría de Vigo after 9 and 12 months, displayed differences to other samples and were located under the lower limit of the 95% confidence interval (AvTD = 94.1% and 94.4%, respectively).

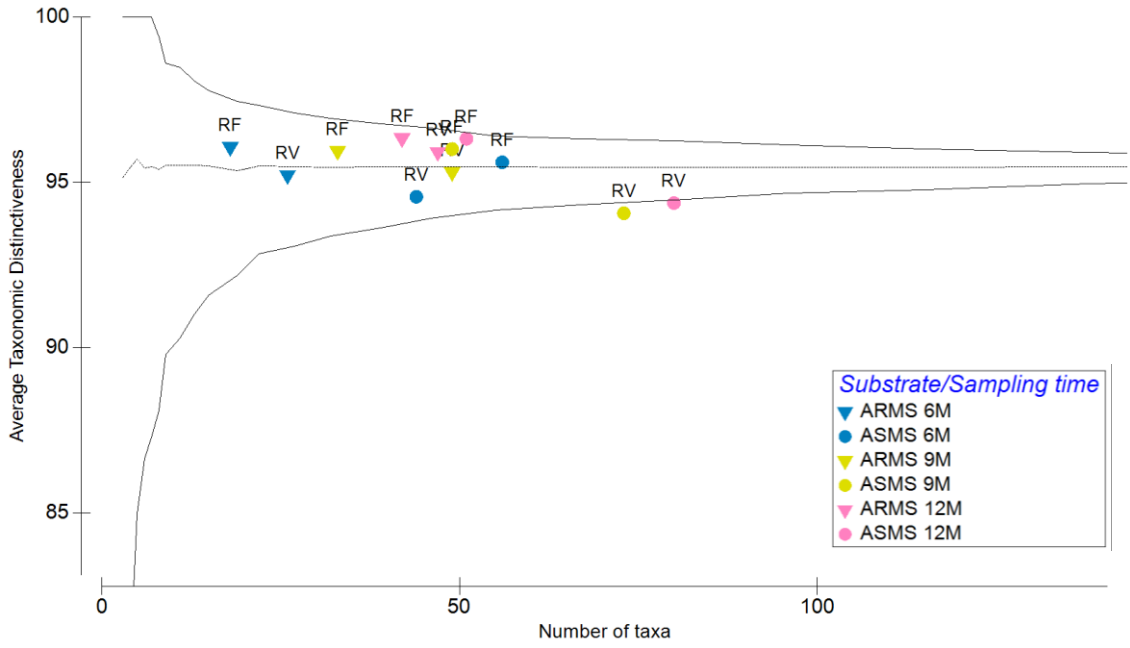


Figure 5.13. Confidence funnel plot of the average taxonomic distinctiveness (AvTD) of the marine macrozoobenthic communities recorded in ARMS and ASMS from Ría de Vigo (RV) and Ría de Ferrol (RF), and across sampling times. Thin line represents theoretical mean (95% confidence interval). 6M – 6 months (blue); 9M – 9 months (yellow); 12M – 12 months (pink).

The non-metric multidimensional scaling, comparing the species detected by both substrates in the two sampling sites, revealed aggregation of the samples in function of the sampling site (Fig.5.14). The substrates also aggregated together, however a higher variation between ARMS is noticed. Furthermore, the obtained dendogram showed a division in three major groups with higher similarity, whose groups aggregated in function of sampling site (except for ARMS from Ferrol after 6 months; Fig S.5.2). Group A comprised ARMS samples from Vigo, after 6 and 9 months of deployment, and were separated from other group - group B, which further sub-divides in group B1, composed by samples from Vigo, and group B2, which consisted of samples from Ferrol. The species which most account for dissimilarity between B1 and

B2 groups (71.02%) were one polychaete (*Bhawania cryptocephala*), two echinoderms (*Antedon bifida* and *Cyllometra manca*) and two hydrozoans (*Halecium beanii* and *Phialella quadrata)* for B1. Furthermore, within and between sampling sites both substrates were very dissimilar (Table S.5.4).

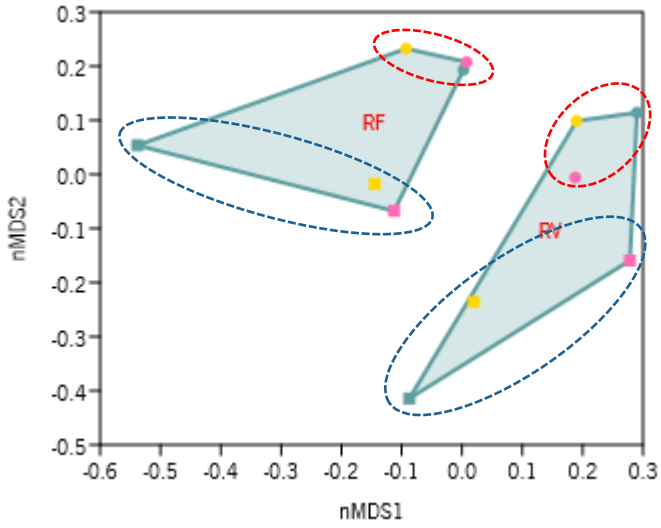


Figure 5.14. Non-metric multidimensional scaling (nMDS), based on Bray-Curtis similarity index, of the species detected in ARMS (■ - green circles) and ASMS (● - red circles). Samples combine the total species detected on each sampling site (RV – Ría de Vigo and RF – Ría de Ferrol) among sampling times (6 months - blue, 9 months - yellow and 12 months - pink).

5.4 Discussion

Our substrate comparison metabarcoding-based study showed the applicability of a standard framework to assess and enable monitoring of coastal macrozoobenthic communities. Artificial substrates of different complexity, namely ARMS and ASMS, demonstrated to be useful as a non-destructive, standard and easy methodology to implement in macroinvertebrates monitoring. However, the results obtained supported different patterns in diversity of macrozoobenthic species detected between substrates, and distinct communities' compositions.

In the current study, a total of 12 marine macrozoobenthic phyla were recovered with high diversity of species. Globally, the communities were dominated by annelids, echinoderms and hydrozoans in terms of number of species. In particular, Hydrozoa was composed by a higher number of species (34 species). Despite being abundant and diverse in benthic ecosystems, and successfully recovered using DNA

metabarcoding (Chapter 4), the high number of species detected could probably be associated to lower genetic distances between congeneric species, and can led misidentified sequences in reference libraries. To obtain certainty of the species-level identifications of hydrozoans, we verified the suitability of the threshold applied, assessing the congeneric genetic distances for both markers (COI and 18S). The majority of the genera recorded (70.6%) displayed a high probability of correct species assignment, and misidentifications explain some of the remaining genus with low interspecific distances (0.1-0.5%).

The results showed low similarity within substrates among sampling times and sites (especially for ARMS – 27.55%). On the other hand, a higher dissimilarity was displayed between ARMS and ASMS, regardless sampling times and study sites. Compared to ARMS, in ASMS we frequently detected more species for each taxonomic group (except for Annelida, Echinodermata and Porifera), and 4 additional taxa were exclusively detected. However, both substrates demonstrated to be complementary in their ability to be colonized by macrozoobenthic species, because a low proportion of species were recorded concurrently in both substrates. The majority of the species occurred exclusively in one substrate, where ASMS not only detected more exclusive species, but also more exclusive taxonomic groups. Previously, Carreira-Flores and collaborators (2020) deployed ASMS in the same location at Ría de Ferrol for 3 and 6 months, and, considering all the species recorded, for the same taxonomic groups we recorded a lower number of species (139 vs 83, respectively). However, they used a different identification methodology (morphology-based), and the sampling was performed in a different year (2018 in Carreira-Flores and 2019 in our study) and in different seasons. Furthermore, they only analyzed mobile fauna, and, if we only compare the mobile fauna, we were able to detected more taxonomic groups. Hence, despite all the uncertainties involved in the comparison of such different studies, according to our previous results (Chapter 4), it appears to be a trend for recording a broader range of taxa with DNA metabarcoding.

Different communities' composition detected between both substrates is likely related to the complexity of the shape of the substrates, particulally due to the complex branching pattern in ASMS, in contrast to the cavities made of bare PVC in ARMS. Additionally, the different tridimensional structure of each substrate was also influenced by exposure to light (even within each ARMS plate; David et al., 2019), to predators and differences in water flow. A fair number of studies have employed ARMS as substrates for species colonization (Ransome et al. 2017; Obst et al., 2020) and recommend it as prime tool for standardized monitoring of macrozoobenthic communities (Obst et al., 2020). However, our results demonstrated that a fair portion of the macrozoobenthos diversity may fail to be captured by this artificial substrate, at least

in the studied region. Hence, the efficiency of this monitoring tool can be partially compromised if there is the risk of systematic overlooking of fractions of diversity. Although missing taxa may not have a great impact in studies aiming for bioassessments of the ecological status, these can be critical for studies aiming for long-term monitoring and assessing global change-induced alterations in species ranges and communities' composition, or detection of NIS. Results obtained for ASMS, in the two study sites and three sampling times, consistently indicate more species and wider taxonomic diversity, as well as 70% different species composition in average. Thus, the complementarity between substrates highlights the need to optimize sampling strategies, where employment of both substrates may be more biodiversity inclusive and detect more species (which are only exclusively detected by one substrate), especially for ASMS.

The communities from each sampling site were well separated in nMDS analysis in two groups, where exclusive species were the drivers for that aggregations. However, within and between sampling sites the analysis demonstrated higher levels of dissimilarity among sampling times. In Ría de Vigo we detected more species and also more taxonomic groups. No pervaniseve species was identified, and only 24.3% and 20.5% of the species detected in RV and RF, respectively, were recorded in all sampling times. In total, as much as 43.1% of the species recorded were exclusive, i.e. only detected in one substrate or sampling time. The overlap of species detected in both substrates was low either in Ría de Vigo (30.3%) or Ría de Ferrol (25%) indicating a high occurrence of exclusive species. Besides consequences of natural sampling variation, exclusive species appear to occur more randomly over time and space suggesting that seasonality may have played a role in the observed patterns. Data from other studies combining ARMS with barcoding revealed patterns similar to our results, with high percentages of exclusive species detected (e.g. 44% of all species, Plaisance et al., 2009, more than 50% of species, Carvalho et al., 2019). Compositional changes of marine macrozoobenthic communities associated with ARMS and ASMS highlighted regional differences, and suggest that a substrate with higher complexity (for example a combination between ARMS and ASMS) will result in more space for species settlement and improve colonization capacity. If possible, the number of replicates should also be increased to attain robust and reliable conclusions about diversity estimates and assessment of regional species dynamics.

Compared to mobile fauna, more species were detected in sessile fauna fractions in both substrates (73.9% ARMS vs 62.8% ASMS). These differences were probably due to the lower number of species commonly recovered by both fractions (36.7%), which is consistent to previous results (Pearman et al., 2020). While annelids and hydrozoans were most abundant in the sessile fraction, crustaceans and echinoderms were

preferentially detected in the mobile fraction. Interestingly, amphipods were exclusively detected in both fractions (e.g. *Caprella acanthifera* only recovered in mobile fraction and *Microdeutopus chelifer* only detected in sessile fraction). Indeed, in mobile fauna samples we recovered species typical of sessile taxa, and vice versa. During sample processing substrates were disassembled, plates in ARMS and branches in ASMS, and some sessile taxa probably broke off and were retained during samples cleaning and sieving. On the other hand, mobile taxa detected in sessile samples may be probably associated to small individuals difficult to separate from the substrate, even using the brush. In each sampling time and in both sampling locations we detected considerably more species in sessile fraction from ASMS than ARMS. However, for three samples of sessile fauna from ARMS we did not have amplification success using COI, which could decrease differences found between sessile fauna from ASMS and ARMS. Previous studies (Cacabelos et al., 2010; Carreira-Flores et al., 2020) referred ARMS as deficient substrates for mobile fauna colonization, since some taxa (e.g. amphipods) are poorly represented. They also detected high levels of mobile fauna colonization in "branched" substrates, such as ASMS. Contrary to these results, in other studies using ARMS authors detected higher diversity in mobile fractions (Carvalho et al., 2019; Leray and Knowlton, 2015). However, they separated the >2 mm fraction from mobile fauna, which could have improved PCR amplification of mobile fauna due to the separation of large-biomass specimens that could have competed disproportionally for amplification.

The results obtained from the AvTD revealed in general high levels of taxonomic diversity for the communities detected on each sample. Two samples, ASMS from Ría de Vigo after 9 and 12 months, were below the confidence limit, indicating a lower phylogenetic complexity in both samples, also visible by the lower degree of taxa diversity displayed in Fig. 5.8. However, these two ASMS were the samples with highest number of species detected (73 and 80, respectively), indicating that higher species richness cannot be directly correlated to high taxonomic diversity. This result was also demonstrated in the relation between number of species and taxonomic diversity, where the sample with highest number of species detected was not the one displaying the highest taxonomic diversity.

The number of species detected varied among sampling times, as well as the representativeness of taxonomic groups. The taxonomic groups showed variations among sampling sites, and only three groups (Annelida, Decapoda and Echinodermata) were detected across all sampling times and substrates. These results are in concordance with the expressive differences obtained between sampling times, highlighting the importance of seasonality in sampling design. Furthermore, the high levels of dissimilarity between

sampling times, including within substrates, demonstrated temporal fluctuations were also dependent on the substrate, which may be a result from interactions between environmental factors and community dynamics. Some species showed time and geographical patterns of occurrence, although this was not evident for a particular taxonomic group. For example, some decapods were firstly detected in Ría de Vigo and in the subsequent sampling time in Ría de Ferrol. The pronounced overtime fluctuations we recorded in species occurrences indicates that sampling only after 12 months of deployment, which is the minimum deployment period normally used in ARMS (Al-Rshaidat et al., 2016; Pearman et al., 2016, 2020; Ransome et al., 2017; Carvalho et al., 2019; Chang et al., 2020; Obst et al., 2020) may fail to capture a fair diversity of taxa and species. Although more or less long periods of deployment may be required for ecological succession to be completed, and for the colonizing assemblage to reach a point of stability mirroring the natural community in that spot, our data indicates that maximum diversity can be reached under 12 months. Although comparisons with other studies are difficult, mostly due to the use of different species identification approaches (e.g. morphology-base identifications) and different sampled locations, a previous study suggested that ASMS complete colonization occurs within 3 months of deployment (Alves, 2017). However, using ASMS in Ría de Vigo we detected more species after 12 months of deployment and after 6 months in Ría de Ferrol (consistent with Carreira-Flores et al., 2020). These different results highlight for the importance of seasonal sampling in long-term monitoring to know when a species is expected to occur, to give information about communities' changes over temporal scales and to signal possible faulty detection of pervasive species, which could flag possible changes on the ecosystem. It should be noted that ARMS monitoring was originally developed for tropical reefs (Zimmerman and Martin, 2004), where ecological succession may take long but, once completed, may be less prone to intense seasonal fluctuations as the ones experienced by the temperate communities such as the ones here studied.

Differences in species detection in both sites and among sampling times could be a result of species distribution patterns. Habitat specific and geographical distribution are the drivers for species trends and patterns. The new complexity of the habitat, a consequence of substrate colonization, as well as new spaces for shelter and settlement can led to shifts in species abundance and occurrence. The heterogeneous community detected in both sampling sites claim for special attention in monitoring studies for the implications on the adopted strategies for biodiversity conservation. Furthermore, the richness in biodiversity detected appears to be the result of high prevalence of exclusive species. Whereas individual species occurrences vary through sampling sites and time, we cannot establish patterns of dominance for such species, since we only have presence/absence data without abundances. However, it is expectable

that less pervasive species will be missed in sampling schemes with low temporal frequency. This creates uncertainties about species richness and distribution, and can provide a deficient characterization of marine communities. Therefore, as these new methodologies involving a combination of artificial substrates and metabarcoding are increasingly implemented in the monitoring of marine macrobenthos, it is expected that a better understanding on the structure and diversity of these communities will be achieved, as well as their species distributions and dynamics.

5.5 Conclusion

In general, ARMS and ASMS promote macrozoobenthic colonization and coupled with DNA metabarcoding revealed as a non-destructive, standard and replicable strategy to capture and monitor coastal communities. The results also illustrated the influence of substrates with different tridimensional structure, which mimic natural habitats and both substrates showed different ability to be colonized by macrozoobenthic species, as well as seasonal variations in the recruitment of zoobenthos. Our results demonstrated that any substrate structure is not able to capture the full diversity of a marine hard-bottom community, especially for ARMS where we detected less taxa. The complementarity recorded between substrates highlighted for the necessity to optimize sampling strategies, where both substrates help to capture their exclusive species. In future monitoring studies, the simultaneous use of both substrates will provide highest efficiency in colonization and will allow to capture a much broader spectrum of taxonomic diversity of coastal ecosystems. Additionally, future research should be devoted to enhance our knowledge on the status of marine ecosystems in the NW Atlantic Iberian coast, especially through the extension of data collection from Portuguese coast. The assessment of such information will allow for comparisons along an extensive and important marine area, improving the potential to obtain a reliable and extensive picture of the situation of the main macrozoobenthic ecosystems and how they respond to environmental changes from local to regional scales.

References

Al-Rshaidat M.M.D., Snider A., Rosebraugh S., Devine A.M., Devine T.D., Plaisance L., Knowlton N., 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. https://doi.org/10.1139/gen-2015-0208.

Alves A.C.M. (2017). Macroalgas formadoras de habitat: Espaço neutro para macrofauna marinha? (Masther's thesis in Ecology). University of Minho, Braga. http://hdl.handle.net/1822/46879.

Barroso C.M., Moreira M.H., Gibbs P.E. (2000). Comparison of imposex and intersex development in four prosobranch species for TBT monitoring of a southern European estuarine system (Ria de Aveiro, NW Portugal). Marine Ecology Progress Series. 201, 221-232. http://dx.doi.org/10.3354/meps201221.

Bianchi C.N., Pronzato R., Cattaneo-Vietti R., Benedetti-Cecchi L., Morri C., Pansini M., Chemello R., Milazzo M., Fraschetti S., Terlizzi A., Peirano A., Salvati E., Benzoni F., Calcinai B., Cerrano C., Bavestrello G. (2004). Hard bottoms. In Gambi M.C., Dappiano M. Mediterranean marine benthos: a manual of methods for its sampling and study. Biologia Marina Mediterranea. 11 (Suppl. 1). SIBM, Genoa.

Borja A., Elliott M., Andersen J.H., Berg T., Carstensen J., Halpern B.S., Heiskanen A.S., Korpinen S., Lowndes J.S.S., Martin G., Rodríguez-Ezpeleta N. (2016). Overview of integrative assessment of marine systems: the ecosystem approach in practice. Frontiers in Marine Science. 3, 20. https://doi.org/10.3389/fmars.2016.00020.

Bourlat S.J., Borja A., Gilbert J., Taylor M.I., Davies N., Weisberg S.B., Griffith J.F., Lettieri T., Field D., Benzie J., Glöckner F.O., Rodríguez-Ezpeleta N., Faith D.P., Obst M. (2013). Genomics in marine monitoring: new opportunities for assessing marine health status. Marine Pollution Bulletin. 74(1), 19-31. http://dx.doi.org/10.1016/j.marpolbul.2013.05.042.

Braukmann T.W.A., Ivanova N.V., Prosser S.W.J., Elbrecht V., Steinke D., Ratnasingham S., de Waard J.R., Sones J.E., Zakharov E.V., Hebert P.D.N. (2019). Metabarcoding a diverse arthropod mock community. Molecular Ecology Resources. 19, 711–727. https://doi. org/10.1111/1755-0998.13008.

Burrows M.T., Schoeman D.S., Buckley L.B., Moore P., Poloczanska E.S., Brander K.M., Brown C., Bruno J.F., Duarte C.M., Halpern B.S., Holding J., Kappel C.V., Kiessling W., O'Connor M.I., Pandolgi J.M., Parmesan C., Scwing F.B., Sydeman W.J., Rchardson A.J. (2011). The pace of shifting climate in marine and terrestrial ecosystems. Science. 334, 652–655. https://doi.org/10.1126/science.1210288.

Cacabelos E., Olabarria C., Incera M., Troncoso J.S. (2010). Effects of habitat structure and tidal height on epifaunal assemblages associated with macroalgae. Estuarine and Coastal Shelf Science. 89, 43–52. http://dx.doi.org/10.1016/j.ecss.2010.05.012.

Cahill A.E., Pearman J.K., Borja A., Carugati L., Carvalho S., Danovaro R., Dashfield S., David R., Féral J., Olenin S., Siaulys A., Somerfield P.J., Trayanova A., Uyarra M.C., Chenuil A. (2018). A comparative analysis of metabarcoding and morphology-based identification of benthic communities across different regional seas. Ecology and Evolution. 8, 8908–8920. https://doi.org/10.1002/ece3.4283.

Carreira-Flores D., Neto R., Ferreira H., Cabecinha E., Díaz-Agras G., Gomes P.T. (2020). Artificial substrates as sampling devices for marine epibenthic fauna: A quest for standardization. Regional Studies in Marine Science. 37, 101331. https://doi.org/10.1016/j.rsma.2020.101331.

Carvalho S., Aylagas E., Villalobos R., Kattan Y., Berumen M., Pearman J.K. (2019). Beyond the visual: using metabarcoding to characterize the hidden reef cryptobiome. Proceedings of the Royal Society B Biological Sciences. 286, 20182697. https://doi.org/10.1098/rspb.2018.2697.

Chang J.J.M., Ip Y.C.A., Bauman A.G., Huang D. (2020). MinION-in-ARMS: nanopore sequencing to expedite barcoding of specimen-rich macrofaunal samples from Autonomous Reef Monitoring Structures. Frontiers in Marine Science. 7, 448. https://doi.org/10.3389/fmars.2020.00448.

Chen H., Boutros P.C. (2011). VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinformatics. 12, 35. http://dx.doi.org/10.1186/1471-2105-12-35.

Clarke K.R., Warwick R.M. (1998). A taxonomic distinctness index and its statistical properties. Journal of Applied Ecology. 35(4), 523-531. https://doi.org/10.1046/j.1365-2664.1998.3540523.x.

Danovaro R. (2003). Pollution threats in the Mediterranean Sea: an overview. Chemistry and Ecology. 19, 15–32. https://doi.org/10.1080/0275754031000081467.

Danovaro R., Carugati L., Berzano M., Cahill A.E., Carvalho S., Chenuil A., Corinaldesi C., Cristina S., David R., Dell'anno A., Dzhembekova N., Garcés E., Gasol J., Goela P., Feral J.P., Ferrera I., Forster R., Kurekin A.A., Rastelli E., Marinova V., Miller P.I., Moncheva S., Newton, A., Pearman J., Pitois S., Renè A., Rodriguez-Ezpeleta N., Saggiomo V., Simis S.G.H., Stefanova K., Wilson C., Lo Martire M., Greco S., Cochrane S., Borja A. (2016). Implementing and innovating marine monitoring approaches for assessing marine environmental status. Frontiers in Marine Science. 3, 213. https://doi.org/10.3389/fmars.2016.00213.

David R., Uyarra M.C., Carvalho S., Anlauf H., Borja A., Cahill A.E., Carugati L., Danovaro R., De Jode A., Féral J., Guillemain D., Lo Martire M., D'Avray L.T.V., Pearman J.K., 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. https://doi.org/10.1016/j.marpolbul.2019.02.066.

Dias J.M.A., Gonzalez R., Garcia C., Diaz-del Rio V. (2002). Sediment distribution patterns on the Galicia-Minho continental shelf. Progress in Oceanography. 52, 215-231. https://doi.org/10.1016/S0079-6611(02)00007-1.

Doney S.C., Ruckelshaus M., Duffy J.E., Barry J.P., Chan F., English C.A., Galindo H.M., Grebmeier J.M., Hollowed A.B., Knowlton N., Polovina J., Rabalais N.N., Sydeman W.J., Talley L.D. (2011). Climate change impacts on marine ecosystems. 4, 11-37. https://doi.org/10.1146/annurev-marine-041911-111611.

Duarte C.M. (2014). Global change and the future ocean: a grand challenge for marine sciences. Frontiers in Marine Science. 1, 63. https://doi.org/10.3389/fmars.2014.00063.

Hammer Ø., Harper D.A.T., Ryan P.D. (2001). PAST: Paleontological statistics software package for education and data analysis. Palaeontologia Electronica. 4(1), 9.

Henriques S., Pais M.P., Batista M.I., Teixeira C.M., Costa M.J., Cabral H. (2014). Can different biological indicators detect similar trends of marine ecosystem degradation? Ecological Indicators. 37, 105-118. https://doi.org/10.1016/j.ecolind.2013.10.017.

Illumina. (2013). 16S Metagenomic sequencing library preparation preparing 16S Ribosomal RNA gene amplicons for the Illumina MiSeq system.

Ivanova N.V., de Waard J.R., Hebert P.D.N. (2006). An inexpensive, automation-friendly protocol for recovering high-quality DNA. Molecular Ecology Notes. 6, 998–1002. https://doi.org/10.1111/j.1471-8286.2006.01428.x.

Kozich J.J., Westcott S.L., Baxter N.T., Highlander S.K., Schloss P.D. (2013). Development of a dualindex sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumine sequencing platform. Applied and Environmental Microbiology. 79, 5112–5120. https://doi.org/10.1128/AEM.01043-13.

Lejzerowicz F., Esling P., Pillet L., Wilding T.A. (2015). High-throughput sequencing and morphology perform equally well for benthic monitoring of marine ecosystems. Scientific Reports. 5, 13932. https://doi.org/10.1038/srep13932.

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. 112(7), 2076–2081. https://doi.org/10.1073/pnas.1424997112.

Leray M., Knowlton N. (2017). Random sampling causes the low reproducibility of rare eukaryotic OTUs in Illumina COI metabarcoding. PeerJ. 5, e3006. https://doi.org/10.7717/peerj.3006.

Leray M., Yang J.Y., Meyer C.P., Mills S.C., Agudelo N., Ranwez V., Boehm J.T., 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. https://doi.org/10.1186/1742-9994-10-34.

Lobo J., Costa P.M., Teixeira M.A.L., Ferreira M.S.G, Costa M.H., Costa F.O. (2013). Enhanced primers for amplification of DNA barcodes from a broad range of marine metazoans. BMC Ecology. 13, 34. https://doi.org/10.1186/1472-6785-13-34.

Lobo J., Shokralla S., Costa M.H., Hajibabaei M., Costa F.O. (2017). DNA metabarcoding for high-throughput monitoring of estuarine macrobenthic communities. Scientific Reports. 7, 15618. https://doi.org/10.1038/s41598-017-15823-6.

Moreira J., Díaz-Agras G., Candás M., Señarís M.P., Urgorri V. (2009). Leptostracans (Crustacea: Phyllocarida) from the Ria de Ferrol (Galicia, NW Iberian Peninsula), with description of a new species of *Nebalia* Leach, 1814. Scientia Marina. 73, 269-285. https://doi.org/0.3989/scimar.2009.73n2269.

Niiranen S., Yletyinen J., Tomczak M.T., Blenckner T., Hjerne O., MacKenzie B.R., Müller-Karulis B., Neumann T., Meier, H. M. (2013). Combined effects of global climate change and regional ecosystem drivers on an exploited marine food web. Global Change Biology. 19(11), 3327-3342. https://doi.org/10.1111/gcb.12309.

Obst M., Exter K., Allcock A.L., Arvanitidis C.D., Bustamante M., Cancio I., Carreira-Flores D., Chatzinikolaou E., Chatzigeorgiou G., Chrismas N., Clark M.S., Comtet T., Dailianis T., Davies N., Deneudt K., de Cerio O.D., Fortic A., Gerovasileiou V., Hablützel P.I., Keklikoglou K., Kotoulas G., Lasota R., Leite B.R., Loisel S., Lévêque L., Levy L., Malachowicz M., Mavric B., Meyer C., Mortelmans J., Norkko J., Pade N., Power A.M., Ramsak A., Reiss H., Solbakken J., Staehr P.A., Sundberg P., Thyrring J., Troncoso J.S., Viard F., Wenne R., Zbawicka M., Pavloudi C. (2020). A marine biodiversity observation network for genetic monitoring of hard-bottom communities (ARMS-MBON). Frontiers in Marine Science. 7. https://doi.org/10.3389/fmars.2020.572680.

Pearman J.K., Anlau, H., Irigoie, 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. https://doi.org/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. https://doi.org/10.1111/mec.15167.

Pearman J.K., Chust G., Aylagas E., Villarino E., Watson J.R., Chenuil A., Borja A., Cahill A.E., Carugati L., Danovaro R., David R., Irigoien X., Mendibil I., Moncheva S., Rodríguez-Ezpeleta N., Uyarra M.C., Carvalho S. (2020). Pan-regional marine benthic cryptobiome biodiversity patterns revealed by metabarcoding Autonomous Reef Monitoring Structures. Molecular Ecology. 29, 4882–4897. https://doi.org/10.1111/mec.15692.

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, 56-65. https://doi.org/10.1016/j.marpolbul.2017.08.001.

Pereira S.G., Lima F.P., Queiroz N.C., Ribeiro P.A., Santos A.M. (2006). Biogeographic patterns of intertidal macroinvertebrates and their association with macroalgae distribution along the Portuguese coast. Hydrobiologia. 555(1), 185. http://dx.doi.org/10.1007/s10750-005-1115-3.

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. https://doi.org/10.1371/journal.pone.0025026.

Plaisance L., Knowlton N., Paulay G., Meyer C. (2009). Reef-associated crustacean fauna: biodiversity estimates using semi-quantitative sampling and DNA barcoding. Coral Reefs. 28(4), 977-986. https://doi.org/10.1007/s00338-009-0543-3.

Prego R., Filgueiras A.V., Santos-Echeandía J. (2008). Temporal and spatial changes of total and labile metal concentration in the surface sediments of the Vigo Ría (NW Iberian Peninsula): influence of anthropogenic sources. Marine Pollution Bulletin. 56, 1031-1042. http://dx.doi.org/10.1016/j.marpolbul.2008.01.036.

Quast C., Pruesse E., Yilmaz P., Gerken J., Schweer T., Yarza P., Peplies J., Glockner F.O. (2013). The SILVA ribossomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research. 41(D1), D590-D596. https://doi.org/10.1093/nar/gks1219.

Ransome E., Geller J.B., Timmers M., Leray M., Mahardini A., Sembiring A., Collins A.G., 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, e0175066. https://doi.org/10.1371/journal.pone.0175066.

Ratnasingham S. (2019). mBRAVE: The Multiplex Barcode Research And Visualization Environment. Biodiversity Information Science and Standards. 3, e37986. https://doi.org/10.3897/biss.3.37986.

Ros M., Vázquez-Luis M., Guerra-García J.M. (2013). The role of marinas and recreational boating in the occurrence and distribution of exotic caprellids (Crustacea: Amphipoda) in the Western Mediterranean: Mallorca Island as a case study. Journal of Sea Research. 83, 94-103. https://doi.org/10.1016/j.seares.2013.04.004.

Rubal M., Veiga P., Vieira R., Sousa-Pinto I. (2011). Seasonal patterns of tidepool macroalgal assemblages in the North of Portugal. Consistence between species and functional group approaches. Journal of Sea Research. 66, 187-194. https://doi.org/10.1016/j.seares.2011.07.003.

Schloss P.D., Westcott S.L., Ryabin T., Hall J.R., Hartmann M., Hollister E.B., Lesniewski R.A., Oakley B.B., Parks D.H., Robinson C.J., Sahl J.W., Stres B., Thallinger G.G., van Horn D.J., Weber C.F. (2009). Introducing mothur:

Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and Environmental Microbiology. 75(23), 7537-7541. https://doi.org/10.1128/AEM.01541-09.

Schmieder R., Edwards R. (2011). Quality control and preprocessing of metagenomic datasets. Bioinformatics. 27(6), 863–4. https://doi.org/10.1093/bioinformatics/btr026.

Sedano F., de Figueroa J.T., Navarro-Barranc C., Ortega E., Guerra-García J.M., Espinosa F. (2020). Do artificial structures cause shifts in epifaunal communities and trophic guilds across different spatial scales? Marine Environmental Research. 104998. https://doi.org/10.1016/j.marenvres.2020.104998.

Shokralla S., Porter T.M., Gibson J.F., Dobosz R., Janzen D.H., Hallwachs W., Golding G.B., Hajibabaei M. (2015). Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform. Scientific Reports. 5, 9687. https://doi.org/10.1038/srep09687.

Spalding M.D., Fox H.E., Gerald G.R., Davidson N., Ferdaña Z.A., Finlayson M., Halper B.S., Jorge M.A., Lombana A.I., Lourie S.A., Martin K.D., McManus E., Molnar J., Recchia C.A., Robertson J. (2007). Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. BioScience. 57, 573–583. http://dx.doi.org/10.1641/B570707.

Steyaert M., Priestley V., Osborne O., Herraiz A., Arnold R., Savolainen V. (2020). Advances in metabarcoding techniques bring us closer to reliable monitoring of the marine benthos. Journal of Applied Ecology. 5(11), 2234-2245. https://doi.org/10.1111/1365-2664.13729.

Stoeck T., Bass D., Nebel M., Christen R., Jones M.D.M., Breiner H.W., Richards T.A. (2010). Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. Molecular Ecology. 19, 21–31. https://doi.org/10.1111/j.1365-294X.2009.04480.x.

Tanhua T., Pouliquen S., Hausman J., O'Brien K., Bricher P., de Bruin T., Buck J.J.H., Burger E.F., Carval T., Casey K.S., Diggs S., Giorgetti A., Glaves H., Harscoat V., Kinkade D., Muelbert J.H., Novellino A., Pfeil B., Pulsifer P.L., van de Putte A., Robinson E., Schaap D., Smirnov A., Smith N., Snowden D., Spears T., Stall S., Tacoma M., Thijsse P., Tronstad S., Vandenberghe T., Wengren M., Wyborn L., Zhao Z.M. (2019). Ocean FAIR Data Services. Frontiers in Marine Science. 6, 440. https://doi.org/10.3389/fmars.2019.00440.

Templado J., Paulay G., Gittenberger A., Meyer C. (2010). Sampling the marine realm. ABC Taxa. 8, 273-307.

Urgorri V., Troncoso J.S., Dobasro J.F. (1992). Malacofauna asociada a una biocenosis de maerl en la ría de Ferrol (Galicia, NO España). Anales de Biología. 18, 161-174.

Warnes G.R., Bolker B., Bonebakker L., Gentleman R., Huber W., Liaw A., Lumley T., Maechler M., Magnusson A., Moeller S., Schwartz M., Venables B. (2020). gplots: Various R programming tools for plotting data. R package version 3.0.3. https://CRAN.R-project.org/package=gplots.

Zimmerman T., Martin J. (2004). Artificial reef matrix structures (Arms): an inexpensive and effective method for collecting coral reef-associated invertebrates. Gulf and Caribbean Research. 16, 59–64. https://doi.org/10.18785/gcr.1601.08.

Chapter 6

Concluding remarks

Through the development and optimization of the sampling and molecular approaches of DNA metabarcoding, this thesis has contributed for the improvement of monitoring protocols for marine macrozoobenthic communities. Additionally, we show that the use of artificial substrates coupled with DNA metabarcoding provides a high-throughput approach able to capture the near-complete profile of species diversity in coastal macrozoobenthic assemblages. This thesis also contributed with novel knowledge of macrozoobenthic species from NW Iberian coast, and supplied insights into patterns of colonization and ecological succession in coastal ecosystems.

The currently available reference libraries for marine macroinvertebrates are fairly incomplete, and partially exhibit low quality records and taxonomic incongruences (Weigand et al., 2019). This affects the accuracy of DNA-based assessments, which may fail in generating a comprehensive taxonomic diagnosis and eventually to provide a reliable profile of the species present in macroinvertebrate communities. One of the main contributions of this thesis was the compilation of a comprehensive reference library of DNA barcodes for the three dominant groups of marine macroinvertebrate species (Annelida, Crustacea and Mollusca) from Atlantic Iberia (Chapter 2; Leite et al., 2020). Combining novel and public data we firstly created a checklist, and then taxonomic discordances were evaluated where most of BINs demonstrated to be taxonomically concordant. However, a relevant portion of the morphospecies were flagged for significant intraspecific divergence, suggesting considerable overlooked diversity. We also performed a gap-analysis, which revealed a high proportion of species still missing DNA barcodes, suggesting the existence of lower levels of completion for marine macroinvertebrate taxa in this region. The reference library provided in Chapter 2 was vital and essential for the correct assignment of species in the following chapters (e.g. to detect possible misidentifications), and have potential utility for further studies (e.g. studies addressing undescribed diversity or cryptic complexes).

Although DNA metabarcoding was demonstrated as a cost-efficient approach, there are still shortfalls associated with the lack of standardization of the protocols (van der Loos and Nijland, 2020). For example, most DNA metabarcoding studies commonly used commercial kits for DNA extraction. Despite their proven efficiency, this is a destructive approach that prevents the posterior verification of the identity of specimens in the case of ambiguous results. In the studies performed through this thesis (Chapters 3-5), we used a non-destructive methodology, which demonstrated to be rapid and cost-effective for application to macrozoobenthic communities. This constitutes an important improvement for molecular identifications and their utility for future research is manifold (e.g. barcoding of new species). A critical decision in a DNA

metabarcoding study is centered on which genomic region should be targeted, since systematic undetected fractions of species diversity may occur in the assessments (Wangensteen et al., 2018). Despite COI has been the most used marker in metabarcoding studies, both COI and 18S have been employed to assess macrozoobenthic diversity through single or, sometimes, multi-locus strategies. Additionally, PCR-based methodologies are highly influenced by amplification biases. Other distinct contribution of this thesis, was the comparison of the ability of COI and 18S to amplify and detect macrozoobenthic taxa (Chapter 3). Furthermore, a combination of primers targeting COI-5P barcode region was also tested. Such an approach has not yet been reported in a similar scale for marine macrozoobenthic species, and, to our best knowledge, despite a few studies used a multi-primer or a multi-locus strategy, but none combined both for marine macrozoobenthos assessment. In our study, although we detected more species with COI than 18S, the species detected varied pronouncedly between primer-pairs, and an extensive complementarity in the species detected by each marker was observed. Only though the employment of both markers, and multiple primer-pairs we assured the highest efficiency for macrozoobenthos detection, with considerably higher number of species and taxonomic groups recovered than using a single marker or primer. Furthermore, only the combination of two primer-pairs targeting COI-5P guaranteed the highest efficiency of species detection, highlighting that no single marker and primer is able to detect the full diversity from a marine macrozoobenthic community. For this reason, we selected this multi-locus and multi-primer strategy to conduct further investigation on macrozoobenthic communities colonizing artificial substrates in NW Iberian coast (Chapter 4 and 5). However, in this thesis we found that primer amplification success is sometimes challenging, and depending on the cost-benefits trade-offs, other primers could be added for monitoring purposes.

Although DNA-based approaches are becoming most efficient, reliable and accurate to assess species diversity, species identification is still commonly performed through morphology-based approaches. However, this is a low-throughput and time-consuming approach, and few studies already compared the ability of both methods to detect macrozoobenthic species (Cahill et al., 2018), and even less used different target marker loci or multi-primer strategy for DNA metabarcoding. Other significant contribution of this thesis, was obtained in Chapter 4, through the combination of morphology and DNA metabarcoding (using the multi-locus and multi-primer strategy designed in Chapter 3) to compare and evaluate the ability of both methods to detect macrozoobenthic species. Compared to morphology, DNA metabarcoding demonstrated higher efficiency and reliability to retrieve more taxa with greater ability to identify at species-level. An important finding was the failure to detect some species exclusively identified through morphology,

which was mainly due to experimental design and sample processing bias. However, incompleteness of reference databases, a problem already discussed in Chapter 3, also explain some of these failures. DNA metabarcoding efficiency offer advantages over morphology, particularly for uncovering some taxonomic groups and small size specimens which may underwent unnoticed when relying only on morphology-based analyses. However, we propose that both methodologies should be used whenever feasible to avoid missing relevant taxa, due to the recognized current limitations of the DNA metabarcoding approaches. Alternatively, a combination of high frequency metabarcoding monitoring intermediated with more spaced morphology-based surveys could provide a means for regular cross-benchmarking the results of both approaches.

Efficient high-throughput sampling strategies aiming to capture the full taxonomic diversity are not well established for marine ecosystems, preventing the implementation of biomonitoring studies in space and time. The common sampling strategies for coastal hard-bottom communities have important shortcomings, since they are destructive and hard to standardize and replicate. Different artificial substrates have been used for different purposes, however, few studies compare their ability to colonize macrozoobenthic species and the possible influence of substrates' material on species colonization (Sedano et al., 2020). Additionally, to our best knowledge, few studies analyzed short- and long-term temporal variation of macrozoobenthic communities colonizing artificial substrates, and few have targeted sampling in locations over a long-time scale.

In Chapter 4, we tested the influence of three substrate materials (Slate, PVC and Granite) on the composition of coastal macrozoobenthic colonization. The substrates were deployed for 3, 7, 10 and 15 months (Ría de Vigo, NW Iberian coast) and morphology-based identification was combined with DNA metabarcoding to assess diversity patterns of macrozoobenthic colonization. Overall, the species detected and taxonomic diversity varied slightly among substrates. Although some taxa revealed intrinsic temporal patterns of colonization, globally, the substrate deployment periods influenced the zoobenthos colonization, reaching the highest number of species after 7 months of deployment. The experimental design combining DNA metabarcoding with artificial substrates allow to evaluate processes of ecological succession of marine macrozoobenthos, discriminating initial phases of colonization, growing through sequential stages without significant differences between substrates. However, in Chapter 4 the impact of the shape of artificial substrates were not evaluated.

Although ARMS have been used and proposed as standard substrates for species colonization (Obst et al., 2020), the taxa preference for this substrate compared to other types and shapes has not been addressed yet. Moreover, to our best knowledge, no studies were found comparing the performance of artificial substrates with different shapes and structures in macrozoobenthic colonization. In Chapter 5, we used DNA metabarcoding to compare overtime macrozoobenthic colonization between tridimensional substrates with different shapes: ARMS and ASMS. Both substrates were deployed in two locations at NW Iberian coast (Ría de Vigo and Ría de Ferrol) and collected after 6, 9 and 12 months to assess species diversity. The two substrates demonstrated complementary in macrozoobenthic colonization, where a large fraction of species was recovered exclusively in a particular substrate (approximately 70%). Both substrates also exhibited different preference to be colonized predominantely by particular taxonomic groups. Relatively to sampling locations, we observed also fair differences between Ría de Vigo and Ría de Ferrol, recording most of the detected species as exclusive to one location. However, using low sampling frequency in monitoring studies is expectable that exclusive or less common species will be missed, creating uncertainties about communities' diversity composition. Therefore, our results demonstrate that the shape of the substrate strongly affected the colonization of species, and that no substrate type is able to capture the full diversity of a marine hard-bottom community. The results also suggest that by using ASMS we enlarged the scope of the phylogenetic diversity recorded, whereas if only one substrate type has been used important fractions of species diversity may underwent overlooked. The different sampling locations used allow to demonstrate that geographic and habitat specificities contribute to determine local macrozoobenthic communities in NW Iberian coast. These characteristics should be taken into account to assess their ecological condition. Moreover, if information on local processes (e.g. physico-chemical parameters) had been explored, more accurate assessments on geographical distribution and successional patterns would be achieved.

This thesis contributed to improve the knowledge of macrozoobenthic diversity in NW Iberian coast, and to design a non-destructive, standard and replicable strategy to capture and monitor coastal zoobenthic communities. However, there are still important shortcomings that need to be addressed before a full transition of monitoring studies to molecular approaches. First, reference libraries should be completed, with comprehensive sampling strategies, ranging from different regions and a broad range of specimens, combined with morphological taxonomy and molecular phylogenetic techniques to produce well represented and quality assured reference libraries. In this thesis we tested a non-destructive methodology for DNA extraction in macrozoobenthic species colonizing artificial substrates, which allowed for the

preservation of specimens after DNA extraction, given the opportunity to develop further barcoding studies, especially for important groups (e.g. *Eumida* complex). Additionally, efforts on morphological identifications coupled with DNA barcoding will help to complete reference databases (e.g. confirm *Elasmopus thalyae* identification, a species still waiting to be "barcoded"). The continuous growth of reference libraries will allow to better understand the diversity of macrozoobenthic species, and to solve discrepancies and incongruences (e.g. hidden diversity flagged in Chapter 3). The use of newly designed bioinformatics tools (Fontes et al., 2020) will allow for quality control and assurance of sequences deposited in databases, providing more accurate assessments in future DNA-based monitoring studies. The developed DNA metabarcoding-based system in this thesis can be further implemented under routine monitoring (e.g. MSFD). Through the combination of ARMS and ASMS with DNA metabarcoding in regular monitoring (short- and long-term) combined with spaced morphology-based assessments (e.g. every 2 years only to confirm some species identifications and to get abundance data) it will be possible to assess more faithfully the impacts and changes in coastal ecosystems. Moreover, because the intensity and frequency of monitoring studies is expected to increase, a better understanding on the structure and diversity of these communities will be achieved, as well as their species distributions and dynamics. Our future research efforts should aim to assess space-time variation of macrozoobenthos with greater detail and accuracy, applying network analysis to gain deeper knowledge on the distribution, variation and functioning of these assemblages. Additionally, it is also important to pursue and intensify sampling in other coastal areas of Iberian Peninsula, particularly extending the surveys to the south along the Portuguese coast, to expand the latitudinal scope and to obtain a broader view of the patterns of variation of these important marine communities.

References

Cahill A.E., Pearman J.K., Borja A., Carugati L., Carvalho S., Danovaro R., Dashfield S., David R., Féral J., Olenin S., Siaulys A., Somerfield P.J., Trayanova A., Uyarra M.C., Chenuil A. (2018). A comparative analysis of metabarcoding and morphology-based identification of benthic communities across different regional seas. Ecology and Evolution. 8, 8908–8920. https://doi.org/10.1002/ece3.4283.

Fontes J.T., Vieira P.E., Ekrem T., Soares P., Costa F.O. (2020). BAGS: an automated Barcode, Audit & Grade System for DNA barcode reference libraries. Molecular Ecology Resources. https://doi.org/10.1111/1755-0998.13262.

Leite B.R., Vieira P.E., Teixeira M.A.L., Lobo-Arteaga J., Hollatz C., Borges L.M.S., Duarte S., Troncoso J.S., Costa F.O. (2020). Gap-analysis and annotated reference library for supporting macroinvertebrate metabarcoding in Atlantic Iberia. Regional Studies in Marine Science. 36, 101307. https://doi.org/10.1016/j.rsma.2020.101307.

Obst M., Exter K., Allcock A.L., Arvanitidis C.D., Bustamante M., Cancio I., Carreira-Flores D., Chatzinikolaou E., Chatzigeorgiou G., Chrismas N., Clark M.S., Comtet T., Dailianis T., Davies N., Deneudt K., de Cerio O.D., Fortic A., Gerovasileiou V., Hablützel P.I., Keklikoglou K., Kotoulas G., Lasota R., Leite B.R., Loisel S., Lévêque L., Levy L., Malachowicz M., Mavric B., Meyer C., Mortelmans J., Norkko J., Pade N., Power A.M., Ramsak A., Reiss H., Solbakken J., Staehr P.A., Sundberg P., Thyrring J., Troncoso J.S., Viard F., Wenne R., Zbawicka M., Pavloudi C. (2020). A marine biodiversity observation network for genetic monitoring of hard-bottom communities (ARMS-MBON). Frontiers in Marine Science. 7. https://doi.org/10.3389/fmars.2020.572680.

Sedano F., Navarro-Barranco C., Guerra-García J.M., Espinosa F. (2020). Understanding the effects of coastal defence structures on marine biota: The role of substrate composition and roughness in structuring sessile, macro-and meiofaunal communities. Marine Pollution Bulletin. 157, 111334. https://doi.org/10.1016/j.marpolbul.2020.111334.

van der Loos L.M., Nijland R. (2020). Biases in bulk: DNA metabarcoding of marine communities and the methodology involved. Molecular Ecology. https://doi.org/10.1111/mec.15592.

Wangensteen O.S., Palacín C., Guardiola M., Turon X. (2018). DNA metabarcoding of littoral hard-bottom communities: high diversity and database gaps revealed by two molecular markers. PeerJ. 6, e4705. http://doi.org/10.7717/peerj.4705.

Weigand H., Beermann A.J., Ciampor F., Costa F.O., Csabai Z., Duarte S., Geiger M.F., Grabowski M., Rimet F., Rulik B., Strand M., Szucsich N., Weigand A.M., Willassen E., Wyler S.A., Bouchez A., Borja A., Ciamporová-Zaťovicová Z., Ferreira S., Dijkstra K.B., Eisendle U., Freyhof J., Gadawski P., Graf W., Haegerbaeumer A., van

der Hoorn B.B., Japoshvili B., Keresztes L., Keskin E., Lesse F., Macher J.N., Mamos T., Paz G., Pesic V., Pfannkuchen D.M., Pfannkuchen M.A., Price B.W., Rinkevich B., Teixeira M.A.L., Várbíró G., Ekrem T. (2019). DNA barcode reference libraries for the monitoring of aquatic biota in Europe: Gap-analysis and recommendations for future work. Science of the Total Environment. 678, 499–524. http://doi.org/10.1016/j.scitotenv.2019.04.247.

Annexes

Annexes of Chapter 3

Table S.3.1. Number of sequences generated in high-throughput sequencing and retaining along processing steps of the bioinformatic pipeline for each primer-pair (COI 1 - mlCOIintF/LoboR1; COI 2 - LCO1490/III_C_R; 18S - TAReuk454FWD1/TAReukREV3) and sampling time. M – Mobile fauna sample; S – Sessile fauna sample.

Sampling time		Primer	Merged reads		Representative sequences		Singletons		Taxonomic assignment	
			M	S	M	S	M	S	M	S
Slate	3 months	COI 1	46821	48487	42919	41357	8	18	33225	18466
		COI 2	28520	44186	24487	38153	5	10	13830	23998
		18S	32103	29163	26855	23918	16	21	4001	5669
	7 months	COI 1	65025	46910	43764	40258	21	12	23835	8824
		COI 2	44884	49676	37822	42905	47	19	22673	18638
		18S	29128	50899	14904	38417	43	78	5454	18387
	10 months	COI 1	50642	41948	44296	37658	10	12	25427	2236
		COI 2	20676	57360	17702	49412	4	6	10556	11591
		18S	30639	51265	24757	28592	57	9	4952	12337
	15 months	COI 1	54441	62908	45614	55693	19	6	40234	47660
		COI 2	81045	34625	69360	28926	19	12	65268	26932
		18S	37460	28246	27192	9047	39	30	11515	6397
PVC	3 months	COI 1	49764	46832	44208	39556	17	0	34845	15292
		COI 2	29012	38421	24743	32524	4	6	22111	22268
		18S	48440	29467	42429	23621	25	39	6283	3798
	7 months	COI 1	57566	56906	49629	46659	19	16	21981	15835
		COI 2	43782	53423	37339	45930	35	29	4978	17166
		18S	31355	41875	13981	27694	55	43	5523	8543
	10 months	COI 1	44944	33326	39098	28571	22	13	19020	5103
		COI 2	35250	40330	29725	34295	6	6	17041	17975
		18S	37447	49078	31439	24700	77	37	4707	6607
	15 months	COI 1	51639	39184	44276	34852	16	4	31085	5602
		COI 2	42191	41608	35974	35214	12	5	29449	16914
		18S	28758	34772	18853	5952	58	38	11686	2018
Granite	3 months	COI 1	-	33622	-	27907	-	3	-	9343
		COI 2	32857	33525	29170	28675	5	8	17762	7827
		18S	39649	44413	35693	37713	3	30	2556	4366
	7 months	COI 1	65025	48298	57083	42669	30	12	22954	11205
		COI 2	24157	43595	20422	37499	15	4	4786	27431
		18S	40919	28071	25105	20614	89	50	6657	9167
	10 months	COI 1	49148	46120	42459	41766	6	4	21800	10102
		COI 2	30344	26092	26018	21997	16	0	15890	12330
		18S	35979	55156	29347	31791	36	33	5185	16099
	15 months	COI 1	37868	33427	33729	29712	15	4	178411	10684
		COI 2	45381	24699	39187	20755	7	1	33272	19523
		18S	32249	33307	20010	12227	29	51	7214	964

Table S.3.2. Taxonomic classification of the marine macroinvertebrates identified at species-level through DNA metabarcoding, with the associated authority based on WoRMS (consulted on 21st July 2020).

Phylum	Class	Order	Family	Genus	Species	Scientific Name	Authority
Annelida	Polychaeta	-	Polygordiidae	*Polygordius*	*appendiculatus*	*Polygordius appendiculatus*	Fraipont, 1887
Annelida	Polychaeta	-	Capitellidae	*Capitella*	*capitata*	*Capitella capitata*	(Fabricius, 1780)
Annelida	Polychaeta	-	Polygordiidae	*Polygordius*	*lacteus*	*Polygordius lacteus*	Schneider, 1868
Annelida	Polychaeta	-	Capitellidae	*Notomastus*	*latericeus*	*Notomastus latericeus*	Sars, 1851
Annelida	Polychaeta	-	Arenicolidae	*Arenicola*	*marina*	*Arenicola marina*	(Linnaeus, 1758)
Annelida	Polychaeta	Amphinomida	Amphinomidae	*Eurythoe*	*complanata*	*Eurythoe complanata*	(Pallas, 1766)
Annelida	Polychaeta	Eunicida	Lumbrineridae	*Scoletoma*	*funchalensis*	*Scoletoma funchalensis*	(Kinberg, 1865)
Annelida	Polychaeta	Eunicida	Eunicidae	*Leodice*	*harassii*	*Leodice harassii*	(Audouin & Milne Edwards, 1833)
Annelida	Polychaeta	Eunicida	Lumbrineridae	*Lumbrineris*	*latreilli*	*Lumbrineris latreilli*	Audouin & Milne Edwards, 1833
Annelida	Polychaeta	Eunicida	Dorvilleidae	*Ophryotrocha*	*puerilis*	*Ophryotrocha puerilis*	Claparède & Mecznikow, 1869
Annelida	Polychaeta	Phyllodocida	Nereididae	*Perinereis*	*aibuhitensis*	*Perinereis aibuhitensis*	(Grube, 1878)
Annelida	Polychaeta	Phyllodocida	Hesionidae	*Syllidia*	*armata*	*Syllidia armata*	Quatrefages, 1866
Annelida	Polychaeta	Phyllodocida	Sigalionidae	*Sthenelais*	*boa*	*Sthenelais boa*	(Johnston, 1833)
Annelida	Polychaeta	Phyllodocida	Syllidae	*Myrianida*	*brachycephala*	*Myrianida brachycephala*	(Marenzeller, 1874)
Annelida	Polychaeta	Phyllodocida	Polynoidae	*Harmothoe*	*clavigera*	*Harmothoe clavigera*	(M. Sars, 1863)

Annelida	Polychaeta	Phyllodocida	Nereididae	*Laeonereis*	*culveri*	*Laeonereis culveri*	(Webster, 1879)
Annelida	Polychaeta	Phyllodocida	Nereididae	*Hediste*	*diversicolor*	*Hediste diversicolor*	(O.F. Müller, 1776)
Annelida	Polychaeta	Phyllodocida	Nereididae	*Platynereis*	*dumerilii*	*Platynereis dumerilii*	(Audouin & Milne Edwards, 1833)
Annelida	Polychaeta	Phyllodocida	Syllidae	*Syllides*	*edentatus*	*Syllides edentatus*	Westheide, 1974
Annelida	Polychaeta	Phyllodocida	Syllidae	*Myrianida*	*edwarsi*	*Myrianida edwarsi*	(Saint Joseph, 1887)
Annelida	Polychaeta	Phyllodocida	Hesionidae	*Psamathe*	*fusca*	*Psamathe fusca*	Johnston, 1836
Annelida	Polychaeta	Phyllodocida	Phyllodocidae	*Sige*	*fusigera*	*Sige fusigera*	Malmgren, 1865
Annelida	Polychaeta	Phyllodocida	Syllidae	*Syllis*	*gracilis*	*Syllis gracilis*	Grube, 1840
Annelida	Polychaeta	Phyllodocida	Phyllodocidae	*Phyllodoce*	*groenlandica*	*Phyllodoce groenlandica*	Örsted, 1842
Annelida	Polychaeta	Phyllodocida	Nephtyidae	*Nephtys*	*hombergii*	*Nephtys hombergii*	Savigny in Lamarck, 1818
Annelida	Polychaeta	Phyllodocida	Polynoidae	*Harmothoe*	*imbricata*	*Harmothoe imbricata*	(Linnaeus, 1767)
Annelida	Polychaeta	Phyllodocida	Polynoidae	*Harmothoe*	*impar*	*Harmothoe impar*	(Johnston, 1839)
Annelida	Polychaeta	Phyllodocida	Nephtyidae	*Nephtys*	*incisa*	*Nephtys incisa*	Malmgren, 1865
Annelida	Polychaeta	Phyllodocida	Pholoidae	*Pholoe*	*inornata*	*Pholoe inornata*	Johnston, 1839
Annelida	Polychaeta	Phyllodocida	Syllidae	*Synmerosyllis*	*lamelligera*	*Synmerosyllis lamelligera*	(Saint-Joseph, 1887)
Annelida	Polychaeta	Phyllodocida	Phyllodocidae	*Eulalia*	*mustela*	*Eulalia mustela*	Pleijel, 1987
Annelida	Polychaeta	Phyllodocida	Syllidae	*Syllis*	*pectinans*	*Syllis pectinans*	Haswell, 1920
Annelida	Polychaeta	Phyllodocida	Nereididae	*Nereis*	*pelagica*	*Nereis pelagica*	Linnaeus, 1758

Annelida	Polychaeta	Phyllodocida	Syllidae	*Myrianida*	*prolifera*	*Myrianida prolifera*	(O.F. Müller, 1788)
Annelida	Polychaeta	Phyllodocida	Phyllodocidae	*Eumida*	*sanguinea*	*Eumida sanguinea*	(Örsted, 1843)
Annelida	Polychaeta	Phyllodocida	Phyllodocidae	*Eulalia*	*viridis*	*Eulalia viridis*	(Linnaeus, 1767)
Annelida	Polychaeta	Phyllodocida	Phyllodocidae	*Paranaitis*	*wahlbergi*	*Paranaitis wahlbergi*	(Malmgren, 1865)
Annelida	Polychaeta	Sabellida	Sabellidae	*Sabella*	*spallanzanii*	*Sabella spallanzanii*	(Gmelin, 1791)
Annelida	Polychaeta	Sabellida	Serpulidae	*Spirobranchus*	*triqueter*	*Spirobranchus triqueter*	(Linnaeus, 1758)
Annelida	Polychaeta	Spionida	Spionidae	*Laonice*	*cirrata*	*Laonice cirrata*	(M. Sars, 1851)
Annelida	Polychaeta	Spionida	Spionidae	*Aonides*	*oxycephala*	*Aonides oxycephala*	(Sars, 1862)
Annelida	Polychaeta	Terebellida	Terebellidae	*Thelepus*	*cincinnatus*	*Thelepus cincinnatus*	(Fabricius, 1780)
Annelida	Polychaeta	Terebellida	Terebellidae	*Amphitritides*	*gracilis*	*Amphitritides gracilis*	(Grube, 1860)
Annelida	Polychaeta	Terebellida	Pectinariidae	*Lagis*	*koreni*	*Lagis koreni*	Malmgren, 1866
Annelida	Polychaeta	Terebellida	Terebellidae	*Terebella*	*lapidaria*	*Terebella lapidaria*	Linnaeus, 1767
Arthropoda	Hexanauplia	Calanoida	Temoridae	*Temora*	*longicornis*	*Temora longicornis*	(Müller O.F., 1785)
Arthropoda	Hexanauplia	Sessilia	Chthamalidae	*Chthamalus*	*montagui*	*Chthamalus montagui*	Southward, 1976
Arthropoda	Hexanauplia	Sessilia	Balanidae	*Perforatus*	*perforatus*	*Perforatus perforatus*	(Bruguière, 1789)
Arthropoda	Hexanauplia	Sessilia	Balanidae	*Balanus*	*trigonus*	*Balanus trigonus*	Darwin, 1854
Arthropoda	Malacostraca	Amphipoda	Caprellidae	*Caprella*	*acanthifera*	*Caprella acanthifera*	Leach, 1814
Arthropoda	Malacostraca	Amphipoda	Aoridae	*Microdeutopus*	*chelifer*	*Microdeutopus chelifer*	(Spence Bate, 1862)

Arthropoda	Malacostraca	Amphipoda	Caprellidae	*Caprella*	*equilibra*	*Caprella equilibra*	Say, 1818
Arthropoda	Malacostraca	Amphipoda	Caprellidae	*Caprella*	*fretensis*	*Caprella fretensis*	Stebbing, 1878
Arthropoda	Malacostraca	Amphipoda	Ischyroceridae	*Jassa*	*herdmani*	*Jassa herdmani*	(Walker, 1893)
Arthropoda	Malacostraca	Amphipoda	Microprotopidae	*Microprotopus*	*maculatus*	*Microprotopus maculatus*	Norman, 1867
Arthropoda	Malacostraca	Amphipoda	Caprellidae	*Phtisica*	*marina*	*Phtisica marina*	Slabber, 1769
Arthropoda	Malacostraca	Amphipoda	Ischyroceridae	*Jassa*	*marmorata*	*Jassa marmorata*	Holmes, 1905
Arthropoda	Malacostraca	Amphipoda	Stenothoidae	*Stenothoe*	*monoculoides*	*Stenothoe monoculoides*	(Montagu, 1813)
Arthropoda	Malacostraca	Amphipoda	Caprellidae	*Caprella*	*penantis*	*Caprella penantis*	Leach, 1814
Arthropoda	Malacostraca	Amphipoda	Maeridae	*Elasmopus*	*rapax*	*Elasmopus rapax*	Costa, 1853
Arthropoda	Malacostraca	Amphipoda	Ampithoidae	*Ampithoe*	*rubricata*	*Ampithoe rubricata*	(Montagu, 1808)
Arthropoda	Malacostraca	Amphipoda	Ischyroceridae	*Jassa*	*slatteryi*	*Jassa slatteryi*	Conlan, 1990
Arthropoda	Malacostraca	Amphipoda	Dexaminidae	*Dexamine*	*spiniventris*	*Dexamine spiniventris*	(Costa, 1853)
Arthropoda	Malacostraca	Amphipoda	Dexaminidae	*Dexamine*	*spinosa*	*Dexamine spinosa*	(Montagu, 1813)
Arthropoda	Malacostraca	Amphipoda	Aoridae	*Aora*	*typica*	*Aora typica*	Krøyer, 1845
Arthropoda	Malacostraca	Amphipoda	Podoceridae	*Podocerus*	*variegatus*	*Podocerus variegatus*	Leach, 1814
Arthropoda	Malacostraca	Decapoda	Thoridae	*Eualus*	*cranchii*	*Eualus cranchii*	(Leach, 1817)
Arthropoda	Malacostraca	Decapoda	Pilumnidae	*Pilumnus*	*hirtellus*	*Pilumnus hirtellus*	(Linnaeus, 1761)
Arthropoda	Malacostraca	Decapoda	Porcellanidae	*Pisidia*	*longicornis*	*Pisidia longicornis*	(Linnaeus, 1767)

Arthropoda	Malacostraca	Decapoda	Carcinidae	*Carcinus*	*maenas*	*Carcinus maenas*	(Linnaeus, 1758)
Arthropoda	Malacostraca	Decapoda	Alpheidae	*Athanas*	*nitescens*	*Athanas nitescens*	(Leach, 1814)
Arthropoda	Malacostraca	Decapoda	Hippolytidae	*Hippolyte*	*varians*	*Hippolyte varians*	Leach, 1814
Arthropoda	Malacostraca	Isopoda	Arcturidae	*Astacilla*	*damnoniensis*	*Astacilla damnoniensis*	(Stebbing, 1874)
Arthropoda	Malacostraca	Isopoda	Sphaeromatidae	*Cymodoce*	*truncata*	*Cymodoce truncata*	Leach, 1814
Arthropoda	Pycnogonida	Pantopoda	Endeidae	*Endeis*	*clipeata*	*Endeis clipeata*	Möbius, 1902
Arthropoda	Pycnogonida	Pantopoda	Ammotheidae	*Achelia*	*echinata*	*Achelia echinata*	Hodge, 1864
Arthropoda	Pycnogonida	Pantopoda	Endeidae	*Endeis*	*spinosa*	*Endeis spinosa*	(Montagu, 1808)
Bryozoa	Gymnolaemata	Cheilostomatida	Aeteidae	*Aetea*	*anguina*	*Aetea anguina*	(Linnaeus, 1758)
Bryozoa	Gymnolaemata	Cheilostomatida	Scruparidae	*Scruparia*	*chelata*	*Scruparia chelata*	(Linnaeus, 1758)
Bryozoa	Gymnolaemata	Cheilostomatida	Microporellidae	*Microporella*	*ciliata*	*Microporella ciliata*	(Pallas, 1766)
Bryozoa	Gymnolaemata	Cheilostomatida	Bugulidae	*Bugulina*	*fulva*	*Bugulina fulva*	(Ryland, 1960)
Bryozoa	Gymnolaemata	Cheilostomatida	Haplopomidae	*Haplopoma*	*graniferum*	*Haplopoma graniferum*	(Johnston, 1847)
Bryozoa	Gymnolaemata	Cheilostomatida	Candidae	*Tricelaria*	*inopinata*	*Tricellaria inopinata*	d'Hondt & Occhipinti Ambrogi, 1985
Bryozoa	Gymnolaemata	Cheilostomatida	Bitectiporidae	*Schizomavella*	*linearis*	*Schizomavella (Schizomavella) linearis*	(Hassall, 1841)
Bryozoa	Gymnolaemata	Cheilostomatida	Bugulidae	*Bugula*	*neritina*	*Bugula neritina*	(Linnaeus, 1758)
Bryozoa	Gymnolaemata	Cheilostomatida	Cryptosulidae	*Cryptosula*	*pallasiana*	*Cryptosula pallasiana*	(Moll, 1803)
Bryozoa	Gymnolaemata	Cheilostomatida	Bugulidae	*Crisularia*	*plumosa*	*Crisularia plumosa*	(Pallas, 1766)

Bryozoa	Gymnolaemata	Cheilostomatida	Escharinidae	*Phaeostachys*	*spinifera*	*Phaeostachys spinifera*	(Johnston, 1847)
Bryozoa	Gymnolaemata	Cheilostomatida	Watersiporidae	*Watersipora*	*subtorquata*	*Watersipora subtorquata*	(d'Orbigny, 1852)
Bryozoa	Gymnolaemata	Cheilostomatida	Bugulidae	*Bugulina*	*turbinata*	*Bugulina turbinata*	(Alder, 1857)
Bryozoa	Gymnolaemata	Ctenostomatida	Vesiculariidae	*Amathia*	*citrina*	*Amathia citrina*	(Hincks, 1877)
Bryozoa	Gymnolaemata	Ctenostomatida	Vesiculariidae	*Amathia*	*gracilis*	*Amathia gracilis*	(Leidy, 1855)
Bryozoa	Gymnolaemata	Ctenostomatida	Vesiculariidae	*Amathia*	*imbricata*	*Amathia imbricata*	(Adams, 1800)
Chordata	Ascidiacea	Stolidobranchia	Styelidae	*Asterocarpa*	*humilis*	*Asterocarpa humilis*	(Heller, 1878)
Cnidaria	Hydrozoa	Anthoathecata	Rathkeidae	*Lizzia*	*blondina*	*Lizzia blondina*	Forbes, 1848
Cnidaria	Hydrozoa	Anthoathecata	Pandeidae	*Amphinema*	*dinema*	*Amphinema dinema*	(Péron & Lesueur, 1810)
Cnidaria	Hydrozoa	Anthoathecata	Bougainvilliidae	*Bougainvillia*	*muscus*	*Bougainvillia muscus*	(Allman, 1863)
Cnidaria	Hydrozoa	Anthoathecata	Eudendriidae	*Eudendrium*	*racemosum*	*Eudendrium racemosum*	(Cavolini, 1785)
Cnidaria	Hydrozoa	Leptothecata	Eirenidae	*Tima*	*bairdii*	*Tima bairdii*	(Johnston, 1833)
Cnidaria	Hydrozoa	Leptothecata	Campanulariidae	*Obelia*	*dichotoma*	*Obelia dichotoma*	(Linnaeus, 1758)
Cnidaria	Hydrozoa	Leptothecata	Campanulariidae	*Obelia*	*geniculata*	*Obelia geniculata*	(Linnaeus, 1758)
Cnidaria	Hydrozoa	Leptothecata	Campanulariidae	*Clytia*	*gracilis*	*Clytia gracilis*	(Sars, 1850)
Cnidaria	Hydrozoa	Leptothecata	Campanulariidae	*Clytia*	*hemisphaerica*	*Clytia hemisphaerica*	(Linnaeus, 1767)
Cnidaria	Hydrozoa	Leptothecata	Campanulariidae	*Campanularia*	*hincksii*	*Campanularia hincksii*	Alder, 1856
Cnidaria	Hydrozoa	Leptothecata	Haleciidae	*Halecium*	*labrosum*	*Halecium labrosum*	Alder, 1859

Cnidaria	Hydrozoa	Leptothecata	Sertularellidae	*Sertularella*	*mediterranea*	*Sertularella mediterranea*	Hartlaub, 1901
Cnidaria	Hydrozoa	Leptothecata	Haleciidae	*Halecium*	*pusillum*	*Halecium pusillum*	Sars, 1856
Cnidaria	Hydrozoa	Leptothecata	Phialellidae	*Phialella*	*quadrata*	*Phialella quadrata*	(Forbes, 1848)
Cnidaria	Hydrozoa	Leptothecata	Plumulariidae	*Plumularia*	*setacea*	*Plumularia setacea*	(Linnaeus, 1758)
Cnidaria	Hydrozoa	Leptothecata	Laodiceidae	*Laodicea*	*undulata*	*Laodicea undulata*	(Forbes & Goodsir, 1853)
Echinodermata	Crinoidea	Comatulida	Antedonidae	*Antedon*	*bifida*	*Antedon bifida*	(Pennant, 1777)
Echinodermata	Echinoidea	Camarodonta	Parechinidae	*Paracentrotus*	*lividus*	*Paracentrotus lividus*	(Lamarck, 1816)
Echinodermata	Echinoidea	Camarodonta	Parechinidae	*Psammechinus*	*miliaris*	*Psammechinus miliaris*	(P.L.S. Müller, 1771)
Echinodermata	Holothuroidea	Dendrochirotida	Cucumariidae	*Aslia*	*lefevrei*	*Aslia lefevrei*	(Barrois, 1882)
Echinodermata	Ophiuroidea	Amphilepidida	Ophiotrichidae	*Ophiothrix*	*fragilis*	*Ophiothrix fragilis*	(Abildgaard in O.F. Müller, 1789)
Mollusca	Bivalvia	Adapedonta	Hiatellidae	*Hiatella*	*arctica*	*Hiatella arctica*	(Linnaeus, 1767)
Mollusca	Bivalvia	Adapedonta	Pharidae	*Ensis*	*ensis*	*Ensis ensis*	(Linnaeus, 1758)
Mollusca	Bivalvia	Cardiida	Cardiidae	*Cerastoderma*	*edule*	*Cerastoderma edule*	(Linnaeus, 1758)
Mollusca	Bivalvia	Cardiida	Tellinidae	*Macomangulus*	*tenuis*	*Macomangulus tenuis*	(da Costa, 1778)
Mollusca	Bivalvia	Limida	Limidae	*Limaria*	*hians*	*Limaria hians*	(Gmelin, 1791)
Mollusca	Bivalvia	Mytilida	Mytilidae	*Musculus*	*discors*	*Musculus discors*	(Linnaeus, 1767)
Mollusca	Bivalvia	Mytilida	Mytilidae	*Mytilus*	*edulis*	*Mytilus edulis*	Linnaeus, 1758
Mollusca	Bivalvia	Mytilida	Mytilidae	*Musculus*	*lateralis*	*Musculus lateralis*	(Say, 1822)

Mollusca	Bivalvia	Pectinida	Anomiidae	*Anomia*	*ephippium*	*Anomia ephippium*	Linnaeus, 1758
Mollusca	Bivalvia	Venerida	Veneridae	*Polititapes*	*aureus*	*Polititapes aureus*	(Gmelin, 1791)
Mollusca	Bivalvia	Venerida	Veneridae	*Ruditapes*	*decussatus*	*Ruditapes decussatus*	(Linnaeus, 1758)
Mollusca	Bivalvia	Venerida	Veneridae	*Ruditapes*	*philippinarum*	*Ruditapes philippinarum*	(A. Adams & Reeve, 1850)
Mollusca	Gastropoda	-	Plakobranchidae	*Elysia*	*viridis*	*Elysia viridis*	(Montagu, 1804)
Mollusca	Gastropoda	-	Patellidae	*Patella*	*vulgata*	*Patella vulgata*	Linnaeus, 1758
Mollusca	Gastropoda	Caenogastropoda	Cerithiopsidae	*Cerithiopsis*	*tubercularis*	*Cerithiopsis tubercularis*	(Montagu, 1803)
Mollusca	Gastropoda	Cephalaspidea	Scaphandridae	*Scaphander*	*lignarius*	*Scaphander lignarius*	(Linnaeus, 1758)
Mollusca	Gastropoda	Littorinimorpha	Calyptraeidae	*Crepipatella*	*dilatata*	*Crepipatella dilatata*	(Lamarck, 1822)
Mollusca	Gastropoda	Littorinimorpha	Rissoidae	*Rissoa*	*lilacina*	*Rissoa lilacina*	Récluz, 1843
Mollusca	Gastropoda	Littorinimorpha	Naticidae	*Euspira*	*nitida*	*Euspira nitida*	(Donovan, 1803)
Mollusca	Gastropoda	Littorinimorpha	Rissoidae	*Rissoa*	*parva*	*Rissoa parva*	(da Costa, 1778)
Mollusca	Gastropoda	Nudibranchia	Facelinidae	*Facelina*	*annulicornis*	*Facelina annulicornis*	(Chamisso & Eysenhardt, 1821)
Mollusca	Gastropoda	Nudibranchia	Facelinidae	*Facelina*	*auriculata*	*Facelina auriculata*	(Müller, 1776)
Mollusca	Gastropoda	Nudibranchia	Facelinidae	*Favorinus*	*branchialis*	*Favorinus branchialis*	(Rathke, 1806)
Mollusca	Gastropoda	Nudibranchia	Trinchesiidae	*Trinchesia*	*caerulea*	*Trinchesia caerulea*	(Montagu, 1804)
Mollusca	Gastropoda	Nudibranchia	Polyceridae	*Limacia*	*clavigera*	*Limacia clavigera*	(O. F. Müller, 1776)
Mollusca	Gastropoda	Nudibranchia	Dotidae	*Doto*	*coronata*	*Doto coronata*	(Gmelin, 1791)

Mollusca	Gastropoda	Nudibranchia	Dotidae	*Doto*	*dunnei*	*Doto dunnei*	Lemche, 1976
Mollusca	Gastropoda	Nudibranchia	Eubranchidae	*Eubranchus*	*exiguus*	*Eubranchus exiguus*	(Alder & Hancock, 1848)
Mollusca	Gastropoda	Nudibranchia	Trinchesiidae	*Trinchesia*	*foliata*	*Trinchesia foliata*	(Forbes & Goodsir, 1839)
Mollusca	Gastropoda	Nudibranchia	Dotidae	*Doto*	*koenneckeri*	*Doto koenneckeri*	Lemche, 1976
Mollusca	Gastropoda	Nudibranchia	Goniodorididae	*Goniodoris*	*nodosa*	*Goniodoris nodosa*	(Montagu, 1808)
Mollusca	Gastropoda	Nudibranchia	Flabellinidae	*Edmundsella*	*pedata*	*Edmundsella pedata*	(Montagu, 1816)
Mollusca	Gastropoda	Nudibranchia	Fionidae	*Fiona*	*pinnata*	*Fiona pinnata*	(Eschscholtz, 1831)
Mollusca	Gastropoda	Nudibranchia	Doridiidae	*Doris*	*pseudoargus*	*Doris pseudoargus*	Rapp, 1827
Mollusca	Gastropoda	Nudibranchia	Polyceridae	*Polycera*	*quadrilineata*	*Polycera quadrilineata*	(O. F. Müller, 1776)
Mollusca	Gastropoda	Nudibranchia	Tergipedidae	*Tergipes*	*tergipes*	*Tergipes tergipes*	(Forsskål in Niebuhr, 1775)
Mollusca	Gastropoda	Nudibranchia	Discodorididae	*Jorunna*	*tomentosa*	*Jorunna tomentosa*	(Cuvier, 1804)
Mollusca	Gastropoda	Pleurobranchida	Pleurobranchidae	*Berthella*	*plumula*	*Berthella plumula*	(Montagu, 1803)
Mollusca	Gastropoda	Trochida	Calliostomatidae	*Calliostoma*	*granulatum*	*Calliostoma granulatum*	(Born, 1778)
Mollusca	Gastropoda	Trochida	Phasianellidae	*Tricolia*	*pullus*	*Tricolia pullus azorica*	(Dautzenberg, 1889)
Mollusca	Gastropoda	Trochida	Phasianellidae	*Tricolia*	*pullus*	*Tricolia pullus canarica*	F. Nordsieck, 1973
Mollusca	Gastropoda	Trochida	Trochidae	*Jujubinus*	*striatus*	*Jujubinus striatus*	(Linnaeus, 1758)
Mollusca	Gastropoda	Trochida	Trochidae	*Steromphala*	*umbilicalis*	*Steromphala umbilicalis*	(da Costa, 1778)
Mollusca	Polyplacophora	Chitonida	Lepidochitonidae	*Lepidochitona*	*cinerea*	*Lepidochitona cinerea*	(Linnaeus, 1767)

Mollusca	Polyplacophora	Chitonida	Acanthochitonidae	*Acanthochitona*	*crinita*	*Acanthochitona crinita*	(Pennant, 1777)
Mollusca	Polyplacophora	Chitonida	Acanthochitonidae	*Acanthochitona*	*fascicularis*	*Acanthochitona fascicularis*	(Linnaeus, 1767)
Mollusca	Polyplacophora	Lepidopleurida	Leptochitonidae	*Lepidopleurus*	*cajetanus*	*Lepidopleurus cajetanus*	(Poli, 1791)
Nemertea	Hoplonemertea	Monostilifera	Tetrastemmatidae	*Tetrastemma*	*candidum*	*Tetrastemma candidum*	(Müller, 1774)
Nemertea	Hoplonemertea	Monostilifera	Tetrastemmatidae	*Tetrastemma*	*coronatum*	*Tetrastemma coronatum*	(Quatrefages, 1846)
Nemertea	Hoplonemertea	Monostilifera	Tetrastemmatidae	*Tetrastemma*	*flavidum*	*Tetrastemma flavidum*	Ehrenberg, 1828
Nemertea	Hoplonemertea	Monostilifera	Emplectonematidae	*Emplectonema*	*gracile*	*Emplectonema gracile*	(Johnston, 1837)
Nemertea	Hoplonemertea	Monostilifera	-	*Vieitezia*	*luzmurubeae*	*Vieitezia luzmurubeae*	Junoy, Andrade & Giribet, 2010
Nemertea	Hoplonemertea	Monostilifera	Tetrastemmatidae	*Tetrastemma*	*vermiculus*	*Tetrastemma vermiculus*	(Quatrefages, 1846)
Nemertea	Palaeonemertea	-	Cephalotrichidae	*Cephalothrix*	*rufifrons*	*Cephalothrix rufifrons*	(Johnston, 1837)
Nemertea	Pilidiophora	Heteronemertea	Lineidae	*Siphonenteron*	*bilineatum*	*Siphonenteron bilineatum*	Meneghini in Renier, 1847
Platyhelminthes	-	Prolecithophora	Plagiostomidae	*Vorticeros*	*auriculatum*	*Vorticeros auriculatum*	(Müller OF, 1784)

Table S.3.3. Genera identified using 18S V4 region for the calculation of genetic distances.

Taxonomic group	Genus
Bryozoa	*Aetea*
Bryozoa	*Amathia*
Bryozoa	*Bugulina*
Bryozoa	*Crisularia*
Bryozoa	*Cryptosula*
Bryozoa	*Haplopoma*
Bryozoa	*Microporella*
Bryozoa	*Obelia*
Bryozoa	*Phaeostachys*
Bryozoa	*Schizomavella*
Bryozoa	*Scruparia*
Bryozoa	*Watersipora*
Crustacea	*Balanus*
Crustacea	*Caprella*
Crustacea	*Carcinus*
Crustacea	*Endeis*
Crustacea	*Eualus*
Crustacea	*Jassa*
Crustacea	*Perforatus*
Crustacea	*Temora*
Echinodermata	*Aslia*
Echinodermata	*Paracentrotus*
Hydrozoa	*Clytia*
Hydrozoa	*Eudendrium*
Hydrozoa	*Halecium*
Hydrozoa	*Laodicea*
Hydrozoa	*Lizzia*
Hydrozoa	*Phialella*
Hydrozoa	*Sertularella*
Hydrozoa	*Tima*
Mollusca	*Anomia*
Mollusca	*Calliostoma*
Mollusca	*Cerastoderma*
Mollusca	*Cuthona*
Mollusca	*Doto*
Mollusca	*Elysia*
Mollusca	*Ensis*
Mollusca	*Fiona*
Mollusca	*Goniodoris*
Mollusca	*Hiatella*
Mollusca	*Jorunna*
Mollusca	*Lepidochitona*
Mollusca	*Lepidopleurus*
Mollusca	*Limaria*
Mollusca	*Macomangulus*
Mollusca	*Musculus*
Mollusca	*Mytilus*
Mollusca	*Ruditapes*
Mollusca	*Scaphander*
Mollusca	*Tergipes*
Nemertea	*Cephalothrix*
Nemertea	*Emplectonema*
Nemertea	*Tetrastemma*
Platyhelminthes	*Vorticeros*
Polychaeta	*Amphitritides*
Polychaeta	*Aonides*
Polychaeta	*Arenicola*
Polychaeta	*Eulalia*
Polychaeta	*Eumida*
Polychaeta	*Eurythoe*
Polychaeta	*Harmothoe*
Polychaeta	*Lumbrineris*
Polychaeta	*Myrianida*
Polychaeta	*Nephtys*
Polychaeta	*Notomastus*
Polychaeta	*Paranaitis*
Polychaeta	*Perinereis*
Polychaeta	*Phyllodoce*
Polychaeta	*Pionosyllis*
Polychaeta	*Platynereis*
Polychaeta	*Polygordius*
Polychaeta	*Sabella*
Polychaeta	*Scoletoma*
Polychaeta	*Sige*
Polychaeta	*Spirobranchus*
Polychaeta	*Syllides*
Polychaeta	*Syllidia*
Polychaeta	*Syllis*
Polychaeta	*Terebella*
Polychaeta	*Thelepus*
Tunicata	*Asterocarpa*

Annexes of Chapter 4

Table S.4.1. Exclusive species detected in one substrate/sampling time combination not included in heatmap analysis.

Taxonomic group	Species
Bryozoa	*Bugulina turbinata*
	Phaeostachys spinifera
Crustacea	*Dexamine spiniventris*
	Caprella fretensis
	Pariambus typicus
	Elasmopus cf. pectenicrus
	Corophium multisetosum
	Atylus swammerdamei
	Podocerus variegatus
	Dynamene edwardsii
	Lekanesphaera rugicauda
	Idotea baltica
	Astacilla damnoniensis
	Carcinus maenas
	Palaemon elegans
	Endeis spinosa
Mollusca	*Anomia ephippium*
	Polititapes aureus
	Limaria hians
	Macomangulus tenuis
	Musculus discors
	Ruditapes philippinarum
	Aequipecten opercularis
	Facelina annulicornis
	Edmundsella pedata
	Berthella plumula
	Doris pseudoargus
	Goniodoris nodosa
	Euspira nitida
	Cingula trifasciata
	Tritia reticulata
	Osilinus lineatus
	Crepipatella dilatata
	Patella vulgata
Echinodermata	*Antedon bifida*
Annelida	*Hediste diversicolor*
	Capitella capitata
	Eulalia viridis
	Psamathe fusca
	Sthenelais boa
	Syllis pectinans
	Sige fusigera
	Nephtys incisa
	Pectinaria koreni
	Polygordius appendiculatus
	Scoletoma funchalensis
	Aonides oxycephala
	Paranaitis wahlbergi
	Thelepus cincinnatus
	Syllides edentatus
Nemertea	*Vieitezia luzmurubeae*
	Siphonenteron bilineatum
	Tetrastemma vermiculus
	Tetrastemma coronatum

Table S.4.2. Taxonomic classification of the marine macroinvertebrates identified at species-level in each substrate through DNA metabarcoding and morphological-based identifications, with the associated authority based on WoRMS (consulted on 21st July 2020). ✓ - with species identification; × - without species identification.

Phylum	Class	Order	Family	Genus	Species	Authority	Morphology	DNA metabarcoding
Annelida	Polychaeta	Amphinomida	Amphinomidae	*Eurythoe*	*complanata*	(Pallas, 1766)	×	✓
Annelida	Polychaeta	Eunicida	Dorvilleidae	*Ophryotrocha*	*puerilis*	Claparède & Mecznikow, 1869	×	✓
Annelida	Polychaeta	Eunicida	Eunicidae	*Leodice*	*harassii*	(Audouin & Milne Edwards, 1833)	×	✓
Annelida	Polychaeta	Eunicida	Lumbrineridae	*Lumbrineris*	*latreilli*	Audouin & Milne Edwards, 1833	×	✓
Annelida	Polychaeta	Eunicida	Lumbrineridae	*Scoletoma*	*funchalensis*	(Kinberg, 1865)	×	✓
Annelida	Polychaeta	Phyllodocida	Hesionidae	*Psamathe*	*fusca*	Johnston, 1836	×	✓
Annelida	Polychaeta	Phyllodocida	Hesionidae	*Syllidia*	*armata*	Quatrefages, 1866	×	✓
Annelida	Polychaeta	Phyllodocida	Nephtyidae	*Nephtys*	*hombergii*	Savigny in Lamarck, 1818	×	✓
Annelida	Polychaeta	Phyllodocida	Nephtyidae	*Nephtys*	*incisa*	Malmgren, 1865	×	✓
Annelida	Polychaeta	Phyllodocida	Nereididae	*Hediste*	*diversicolor*	(O.F. Müller, 1776)	×	✓
Annelida	Polychaeta	Phyllodocida	Nereididae	*Laeonereis*	*culveri*	(Webster, 1879)	×	✓
Annelida	Polychaeta	Phyllodocida	Nereididae	*Nereis*	*pelagica*	Linnaeus, 1758	✓	✓
Annelida	Polychaeta	Phyllodocida	Nereididae	*Perinereis*	*aibuhitensis*	(Grube, 1878)	×	✓
Annelida	Polychaeta	Phyllodocida	Nereididae	*Platynereis*	*dumerilii*	(Audouin & Milne Edwards, 1833)	✓	✓
Annelida	Polychaeta	Phyllodocida	Pholoidae	*Pholoe*	*inornata*	Johnston, 1839	×	✓
Annelida	Polychaeta	Phyllodocida	Phyllodocidae	*Eulalia*	*mustela*	Pleijel, 1987	×	✓
Annelida	Polychaeta	Phyllodocida	Phyllodocidae	*Eulalia*	*viridis*	(Linnaeus, 1767)	×	✓
Annelida	Polychaeta	Phyllodocida	Phyllodocidae	*Eumida*	*sanguinea*	(Örsted, 1843)	✓	✓
Annelida	Polychaeta	Phyllodocida	Phyllodocidae	*Paranaitis*	*wahlbergi*	(Malmgren, 1865)	×	✓
Annelida	Polychaeta	Phyllodocida	Phyllodocidae	*Phyllodoce*	*groenlandica*	Örsted, 1842	×	✓
Annelida	Polychaeta	Phyllodocida	Phyllodocidae	*Sige*	*fusigera*	Malmgren, 1865	×	✓
Annelida	Polychaeta	Phyllodocida	Polynoidae	*Harmothoe*	*clavigera*	(M. Sars, 1863)	×	✓

Annelida	Polychaeta	Phyllodocida	Polynoidae	*Harmothoe*	*imbricata*	(Linnaeus, 1767)	×	✓
Annelida	Polychaeta	Phyllodocida	Polynoidae	*Harmothoe*	*impar*	(Johnston, 1839)	✓	✓
Annelida	Polychaeta	Phyllodocida	Polynoidae	*Lepidonotus*	*squamatus*	(Linnaeus, 1758)	✓	×
Annelida	Polychaeta	Phyllodocida	Sigalionidae	*Sthenelais*	*boa*	(Johnston, 1833)	×	✓
Annelida	Polychaeta	Phyllodocida	Syllidae	*Myrianida*	*brachycephala*	(Marenzeller, 1874)	×	✓
Annelida	Polychaeta	Phyllodocida	Syllidae	*Myrianida*	*edwarsi*	(Saint Joseph, 1887)	×	✓
Annelida	Polychaeta	Phyllodocida	Syllidae	*Myrianida*	*prolifera*	(O.F. Müller, 1788)	×	✓
Annelida	Polychaeta	Phyllodocida	Syllidae	*Pionosyllis*	*lamelligera*	(Saint-Joseph, 1887)	×	✓
Annelida	Polychaeta	Phyllodocida	Syllidae	*Syllides*	*edentatus*	Westheide, 1974	×	✓
Annelida	Polychaeta	Phyllodocida	Syllidae	*Syllis*	*gracilis*	Grube, 1840	×	✓
Annelida	Polychaeta	Phyllodocida	Syllidae	*Syllis*	*pectinans*	Haswell, 1920	×	✓
Annelida	Polychaeta	Phyllodocida	Syllidae	*Trypanosyllis*	*zebra*	(Grube, 1860)	✓	×
Annelida	Polychaeta	Sabellida	Sabellidae	*Bispira*	*crassicornis*	(Sars, 1851)	×	✓
Annelida	Polychaeta	Sabellida	Sabellidae	*Sabella*	*spallanzanii*	(Gmelin, 1791)	×	✓
Annelida	Polychaeta	Sabellida	Serpulidae	*Spirobranchus*	*triqueter*	(Linnaeus, 1758)	✓	✓
Annelida	Polychaeta	Spionida	Spionidae	*Aonides*	*oxycephala*	(Sars, 1862)	×	✓
Annelida	Polychaeta	Spionida	Spionidae	*Laonice*	*cirrata*	(M. Sars, 1851)	×	✓
Annelida	Polychaeta	Spionida	Spionidae	*Prionospio*	*fallax*	Söderström, 1920	×	✓
Annelida	Polychaeta	Terebellida	Pectinariidae	*Lagis*	*koreni*	Malmgren, 1866	×	✓
Annelida	Polychaeta	Terebellida	Terebellidae	*Amphitritides*	*gracilis*	(Grube, 1860)	×	✓
Annelida	Polychaeta	Terebellida	Terebellidae	*Terebella*	*lapidaria*	Linnaeus, 1767	×	✓
Annelida	Polychaeta	Terebellida	Terebellidae	*Thelepus*	*cincinnatus*	(Fabricius, 1780)	×	✓
Annelida	Polychaeta	-	Arenicolidae	*Arenicola*	*marina*	(Linnaeus, 1758)	×	✓
Annelida	Polychaeta	-	Capitellidae	*Capitella*	*capitata*	(Fabricius, 1780)	×	✓
Annelida	Polychaeta	-	Capitellidae	*Notomastus*	*latericeus*	Sars, 1851	×	✓
Annelida	Polychaeta	-	Polygordiidae	*Polygordius*	*appendiculatus*	Fraipont, 1887	×	✓
Annelida	Polychaeta	-	Polygordiidae	*Polygordius*	*lacteus*	Schneider, 1868	×	✓

Arthropoda	Hexanauplia	Calanoida	Temoridae	*Temora*	*longicornis*	(Müller O.F., 1785)	×	✓
Arthropoda	Hexanauplia	Harpacticoida	Thalestridae	*Parathalestris*	*harpactoides* cf.	(Claus, 1863)	✓	×
Arthropoda	Hexanauplia	Sessilia	Balanidae	*Balanus*	*trigonus*	Darwin, 1854	×	✓
Arthropoda	Hexanauplia	Sessilia	Balanidae	*Perforatus*	*perforatus*	(Bruguière, 1789)	✓	✓
Arthropoda	Hexanauplia	Sessilia	Chthamalidae	*Chthamalus*	*montagui*	Southward, 1976	×	✓
Arthropoda	Malacostraca	Amphipoda	Ampithoidae	*Ampithoe*	*rubricata*	(Montagu, 1808)	✓	✓
Arthropoda	Malacostraca	Amphipoda	Aoridae	*Aora*	*typica*	Krøyer, 1845	×	✓
Arthropoda	Malacostraca	Amphipoda	Aoridae	*Microdeutopus*	*chelifer*	(Spence Bate, 1862)	×	✓
Arthropoda	Malacostraca	Amphipoda	Atylidae	*Nototropis*	*swammerdamei* cf.	(H. Milne Edwards, 1830)	✓	×
Arthropoda	Malacostraca	Amphipoda	Caprellidae	*Caprella*	*acanthifera*	Leach, 1814	✓	✓
Arthropoda	Malacostraca	Amphipoda	Caprellidae	*Caprella*	*danilevskii*	Czerniavski, 1868	✓	×
Arthropoda	Malacostraca	Amphipoda	Caprellidae	*Caprella*	*equilibra*	Say, 1818	✓	✓
Arthropoda	Malacostraca	Amphipoda	Caprellidae	*Caprella*	*fretensis*	Stebbing, 1878	×	✓
Arthropoda	Malacostraca	Amphipoda	Caprellidae	*Caprella*	*liparotensis*	Haller, 1879	✓	×
Arthropoda	Malacostraca	Amphipoda	Caprellidae	*Caprella*	*penantis*	Leach, 1814	✓	✓
Arthropoda	Malacostraca	Amphipoda	Caprellidae	*Pariambus*	*typicus*	(Krøyer, 1845)	✓	×
Arthropoda	Malacostraca	Amphipoda	Caprellidae	*Phtisica*	*marina*	Slabber, 1769	✓	✓
Arthropoda	Malacostraca	Amphipoda	Caprellidae	*Pseudoprotella*	*phasma*	(Montagu, 1804)	✓	×
Arthropoda	Malacostraca	Amphipoda	Corophiidae	*Corophium*	*multisetosum* cf.	Stock, 1952	✓	×
Arthropoda	Malacostraca	Amphipoda	Dexaminidae	*Dexamine*	*spiniventris*	(Costa, 1853)	×	✓
Arthropoda	Malacostraca	Amphipoda	Dexaminidae	*Dexamine*	*spinosa*	(Montagu, 1813)	×	✓
Arthropoda	Malacostraca	Amphipoda	Ischyroceridae	*Jassa*	*herdmani*	(Walker, 1893)	✓	✓
Arthropoda	Malacostraca	Amphipoda	Ischyroceridae	*Jassa*	*marmorata*	Holmes, 1905	✓	✓
Arthropoda	Malacostraca	Amphipoda	Ischyroceridae	*Jassa*	*slatteryi*	Conlan, 1990	✓	✓
Arthropoda	Malacostraca	Amphipoda	Maeridae	*Elasmopus*	*brasiliensis* cf.	(Dana, 1853)	✓	×
Arthropoda	Malacostraca	Amphipoda	Maeridae	*Elasmopus*	*pectenicrus*	(Spence Bate, 1862)	✓	×
Arthropoda	Malacostraca	Amphipoda	Maeridae	*Elasmopus*	*rapax*	Costa, 1853	✓	✓

Arthropoda	Malacostraca	Amphipoda	Maeridae	*Elasmopus*	*thalyae*	Gouillieux & Sorbe, 2015	✓	×
Arthropoda	Malacostraca	Amphipoda	Melitidae	*Melita*	*palmata* cf.	(Montagu, 1804)	✓	×
Arthropoda	Malacostraca	Amphipoda	Microprotopidae	*Microprotopus*	*maculatus*	Norman, 1867	✓	✓
Arthropoda	Malacostraca	Amphipoda	Photidae	*Gammaropsis*	*maculata* cf.	(Johnston, 1828)	✓	×
Arthropoda	Malacostraca	Amphipoda	Podoceridae	*Podocerus*	*variegatus*	Leach, 1814	×	✓
Arthropoda	Malacostraca	Amphipoda	Stenothoidae	*Stenothoe*	*monoculoides*	(Montagu, 1813)	✓	✓
Arthropoda	Malacostraca	Decapoda	Alpheidae	*Athanas*	*nitescens*	(Leach, 1814 [in Leach, 1813-1815])	✓	✓
Arthropoda	Malacostraca	Decapoda	Carcinidae	*Carcinus*	*maenas*	(Linnaeus, 1758)	×	✓
Arthropoda	Malacostraca	Decapoda	Hippolytidae	*Hippolyte*	*varians*	Leach, 1814 [in Leach, 1813-1815]	✓	✓
Arthropoda	Malacostraca	Decapoda	Palaemonidae	*Palaemon*	*elegans*	Rathke, 1837	✓	×
Arthropoda	Malacostraca	Decapoda	Pilumnidae	*Pilumnus*	*hirtellus*	(Linnaeus, 1761)	✓	✓
Arthropoda	Malacostraca	Decapoda	Porcellanidae	*Pisidia*	*longicornis*	(Linnaeus, 1767)	✓	✓
Arthropoda	Malacostraca	Decapoda	Thoridae	*Eualus*	*cranchii*	(Leach, 1817 [in Leach, 1815-1875])	✓	✓
Arthropoda	Malacostraca	Decapoda	Xanthidae	*Xantho*	*pilipes*	A. Milne-Edwards, 1867	✓	×
Arthropoda	Malacostraca	Isopoda	Anthuridae	*Cyathura*	*carinata*	(Krøyer, 1847)	✓	×
Arthropoda	Malacostraca	Isopoda	Arcturidae	*Astacilla*	*damnoniensis*	(Stebbing, 1874)	×	✓
Arthropoda	Malacostraca	Isopoda	Idoteidae	*Idotea*	*baltica*	(Pallas, 1772)	✓	×
Arthropoda	Malacostraca	Isopoda	Sphaeromatidae	*Cymodoce*	*truncata*	Leach, 1814	✓	✓
Arthropoda	Malacostraca	Isopoda	Sphaeromatidae	*Dynamene*	*bidentata*	(Adams, 1800)	✓	×
Arthropoda	Malacostraca	Isopoda	Sphaeromatidae	*Dynamene*	*edwardsi*	(Lucas, 1849)	✓	×
Arthropoda	Malacostraca	Isopoda	Sphaeromatidae	*Dynamene*	*magnitorata*	Holdich, 1968	✓	×
Arthropoda	Malacostraca	Isopoda	Sphaeromatidae	*Lekanesphaera*	*rugicauda*	(Leach, 1814)	✓	×
Arthropoda	Pycnogonida	Pantopoda	Ammotheidae	*Achelia*	*echinata*	Hodge, 1864	✓	✓
Arthropoda	Pycnogonida	Pantopoda	Endeidae	*Endeis*	*clipeata*	Möbius, 1902	×	✓
Arthropoda	Pycnogonida	Pantopoda	Endeidae	*Endeis*	*spinosa*	(Montagu, 1808)	×	✓
Bryozoa	Gymnolaemata	Cheilostomatida	Aeteidae	*Aetea*	*anguina*	(Linnaeus, 1758)	×	✓
Bryozoa	Gymnolaemata	Cheilostomatida	Bitectiporidae	*Schizomavella*	*linearis*	(Hassall, 1841)	×	✓

Bryozoa	Gymnolaemata	Cheilostomatida	Bugulidae	*Bugula*	*neritina*	(Linnaeus, 1758)	×	✓
Bryozoa	Gymnolaemata	Cheilostomatida	Bugulidae	*Bugulina*	*fulva*	(Ryland, 1960)	×	✓
Bryozoa	Gymnolaemata	Cheilostomatida	Bugulidae	*Bugulina*	*turbinata*	(Alder, 1857)	×	✓
Bryozoa	Gymnolaemata	Cheilostomatida	Bugulidae	*Crisularia*	*plumosa*	(Pallas, 1766)	×	✓
Bryozoa	Gymnolaemata	Cheilostomatida	Candidae	*Tricelaria*	*inopinata*	d'Hondt & Occhipinti Ambrogi, 1985	✓	✓
Bryozoa	Gymnolaemata	Cheilostomatida	Cryptosulidae	*Cryptosula*	*pallasiana*	(Moll, 1803)	×	✓
Bryozoa	Gymnolaemata	Cheilostomatida	Escharinidae	*Phaeostachys*	*spinifera*	(Johnston, 1847)	×	✓
Bryozoa	Gymnolaemata	Cheilostomatida	Haplopomidae	*Haplopoma*	*graniferum*	(Johnston, 1847)	×	✓
Bryozoa	Gymnolaemata	Cheilostomatida	Microporellidae	*Microporella*	*ciliata*	(Pallas, 1766)	×	✓
Bryozoa	Gymnolaemata	Cheilostomatida	Scruparіidae	*Scruparia*	*chelata*	(Linnaeus, 1758)	×	✓
Bryozoa	Gymnolaemata	Cheilostomatida	Watersiporidae	*Watersipora*	*subtorquata*	(d'Orbigny, 1852)	✓	✓
Bryozoa	Gymnolaemata	Ctenostomatida	Vesiculariidae	*Amathia*	*citrina*	(Hincks, 1877)	×	✓
Bryozoa	Gymnolaemata	Ctenostomatida	Vesiculariidae	*Amathia*	*gracilis*	(Leidy, 1855)	×	✓
Bryozoa	Gymnolaemata	Ctenostomatida	Vesiculariidae	*Amathia*	*imbricata*	(Adams, 1800)	×	✓
Chordata	Ascidiacea	Stolidobranchia	Styelidae	*Asterocarpa*	*humilis*	(Heller, 1878)	✓	✓
Cnidaria	Hydrozoa	Anthoathecata	Bougainvilliidae	*Bougainvillia*	*muscus*	(Allman, 1863)	✓	✓
Cnidaria	Hydrozoa	Anthoathecata	Eudendriidae	*Eudendrium*	*racemosum*	(Cavolini, 1785)	×	✓
Cnidaria	Hydrozoa	Anthoathecata	Pandeidae	*Amphinema*	*dinema*	(Péron & Lesueur, 1810)	×	✓
Cnidaria	Hydrozoa	Anthoathecata	Rathkeidae	*Lizzia*	*blondina*	Forbes, 1848	×	✓
Cnidaria	Hydrozoa	Leptothecata	Campanulariidae	*Campanularia*	*hincksii*	Alder, 1856	×	✓
Cnidaria	Hydrozoa	Leptothecata	Campanulariidae	*Clytia*	*gracilis*	(Sars, 1850)	×	✓
Cnidaria	Hydrozoa	Leptothecata	Campanulariidae	*Clytia*	*hemisphaerica*	(Linnaeus, 1767)	×	✓
Cnidaria	Hydrozoa	Leptothecata	Campanulariidae	*Obelia*	*dichotoma*	(Linnaeus, 1758)	×	✓
Cnidaria	Hydrozoa	Leptothecata	Campanulariidae	*Obelia*	*geniculata*	(Linnaeus, 1758)	×	✓
Cnidaria	Hydrozoa	Leptothecata	Eirenidae	*Tima*	*bairdii*	(Johnston, 1833)	×	✓
Cnidaria	Hydrozoa	Leptothecata	Haleciidae	*Halecium*	*labrosum*	Alder, 1859	×	✓
Cnidaria	Hydrozoa	Leptothecata	Haleciidae	*Halecium*	*pusillum*	Sars, 1856	×	✓

Cnidaria	Hydrozoa	Leptothecata	Laodiceidae	*Laodicea*	*undulata*	(Forbes & Goodsir, 1853)	×	✓
Cnidaria	Hydrozoa	Leptothecata	Phialellidae	*Phialella*	*quadrata*	(Forbes, 1848)	×	✓
Cnidaria	Hydrozoa	Leptothecata	Plumulariidae	*Plumularia*	*setacea*	(Linnaeus, 1758)	✓	✓
Cnidaria	Hydrozoa	Leptothecata	Sertularellidae	*Sertularella*	*mediterranea*	Hartlaub, 1901	×	✓
Echinodermata	Crinoidea	Comatulida	Antedonidae	*Antedon*	*bifida*	(Pennant, 1777)	×	✓
Echinodermata	Echinoidea	Camarodonta	Parechinidae	*Paracentrotus*	*lividus*	(Lamarck, 1816)	✓	✓
Echinodermata	Echinoidea	Camarodonta	Parechinidae	*Psammechinus*	*miliaris*	(P.L.S. Müller, 1771)	✓	✓
Echinodermata	Holothuroidea	Dendrochirotida	Cucumariidae	*Aslia*	*lefevrei*	(Barrois, 1882)	×	✓
Echinodermata	Ophiuroidea	Amphilepidida	Ophiotrichidae	*Ophiothrix*	*fragilis*	(Abildgaard in O.F. Müller, 1789)	×	✓
Mollusca	Bivalvia	Adapedonta	Hiatellidae	*Hiatella*	*arctica*	(Linnaeus, 1767)	✓	✓
Mollusca	Bivalvia	Adapedonta	Pharidae	*Ensis*	*ensis*	(Linnaeus, 1758)	×	✓
Mollusca	Bivalvia	Cardiida	Cardiidae	*Cerastoderma*	*edule*	(Linnaeus, 1758)	×	✓
Mollusca	Bivalvia	Cardiida	Tellinidae	*Macomangulus*	*tenuis*	(da Costa, 1778)	×	✓
Mollusca	Bivalvia	Limida	Limidae	*Limaria*	*hians*	(Gmelin, 1791)	×	✓
Mollusca	Bivalvia	Mytilida	Mytilidae	*Musculus*	*discors*	(Linnaeus, 1767)	×	✓
Mollusca	Bivalvia	Mytilida	Mytilidae	*Musculus*	*lateralis*	(Say, 1822)	×	✓
Mollusca	Bivalvia	Mytilida	Mytilidae	*Mytilus*	*edulis*	Linnaeus, 1758	×	✓
Mollusca	Bivalvia	Mytilida	Mytilidae	*Mytilus*	*galloprovincialis*	Lamarck, 1819	✓	×
Mollusca	Bivalvia	Pectinida	Anomiidae	*Anomia*	*ephippium*	Linnaeus, 1758	×	✓
Mollusca	Bivalvia	Pectinida	Pectinidae	*Aequipecten*	*opercularis* cf.	(Linnaeus, 1758)	✓	×
Mollusca	Bivalvia	Pectinida	Pectinidae	*Pecten*	*maximus*	(Linnaeus, 1758)	✓	×
Mollusca	Bivalvia	Venerida	Veneridae	*Polititapes*	*aureus*	(Gmelin, 1791)	×	✓
Mollusca	Bivalvia	Venerida	Veneridae	*Ruditapes*	*decussatus*	(Linnaeus, 1758)	✓	✓
Mollusca	Bivalvia	Venerida	Veneridae	*Ruditapes*	*philippinarum*	(A. Adams & Reeve, 1850)	×	✓
Mollusca	Gastropoda	Caenogastropoda	Cerithiopsidae	*Cerithiopsis*	*tubercularis*	(Montagu, 1803)	✓	✓
Mollusca	Gastropoda	Cephalaspidea	Scaphandridae	*Scaphander*	*lignarius*	(Linnaeus, 1758)	×	✓
Mollusca	Gastropoda	Littorinimorpha	Calyptraeidae	*Crepipatella*	*dilatata*	(Lamarck, 1822)	×	✓

Mollusca	Gastropoda	Littorinimorpha	Littorinidae	*Lacuna*	*pallidula* cf.	(da Costa, 1778)	✓	×
Mollusca	Gastropoda	Littorinimorpha	Littorinidae	*Littorina*	*saxatilis* cf.	(Olivi, 1792)	✓	×
Mollusca	Gastropoda	Littorinimorpha	Naticidae	*Euspira*	*nitida*	(Donovan, 1803)	×	✓
Mollusca	Gastropoda	Littorinimorpha	Rissoidae	*Cingula*	*trifasciata*	(J. Adams, 1800)	✓	×
Mollusca	Gastropoda	Littorinimorpha	Rissoidae	*Rissoa*	*lilacina*	Récluz, 1843	✓	✓
Mollusca	Gastropoda	Littorinimorpha	Rissoidae	*Rissoa*	*parva*	(da Costa, 1778)	✓	✓
Mollusca	Gastropoda	Neogastropoda	Nassariidae	*Tritia*	*incrassata*	(Strøm, 1768)	✓	×
Mollusca	Gastropoda	Neogastropoda	Nassariidae	*Tritia*	*reticulata*	(Linnaeus, 1758)	✓	×
Mollusca	Gastropoda	Nudibranchia	Discodorididae	*Jorunna*	*tomentosa*	(Cuvier, 1804)	×	✓
Mollusca	Gastropoda	Nudibranchia	Doridiidae	*Doris*	*pseudoargus*	Rapp, 1827	×	✓
Mollusca	Gastropoda	Nudibranchia	Dotidae	*Doto*	*coronata*	(Gmelin, 1791)	✓	✓
Mollusca	Gastropoda	Nudibranchia	Dotidae	*Doto*	*dunnei*	Lemche, 1976	×	✓
Mollusca	Gastropoda	Nudibranchia	Dotidae	*Doto*	*koenneckeri*	Lemche, 1976	×	✓
Mollusca	Gastropoda	Nudibranchia	Eubranchidae	*Eubranchus*	*exiguus*	(Alder & Hancock, 1848)	×	✓
Mollusca	Gastropoda	Nudibranchia	Facelinidae	*Facelina*	*annulicornis*	(Chamisso & Eysenhardt, 1821)	×	✓
Mollusca	Gastropoda	Nudibranchia	Facelinidae	*Facelina*	*auriculata*	(Müller, 1776)	×	✓
Mollusca	Gastropoda	Nudibranchia	Facelinidae	*Favorinus*	*branchialis*	(Rathke, 1806)	×	✓
Mollusca	Gastropoda	Nudibranchia	Fionidae	*Fiona*	*pinnata*	(Eschscholtz, 1831)	×	✓
Mollusca	Gastropoda	Nudibranchia	Flabellinidae	*Edmundsella*	*pedata*	(Montagu, 1816)	×	✓
Mollusca	Gastropoda	Nudibranchia	Goniodorididae	*Goniodoris*	*nodosa*	(Montagu, 1808)	×	✓
Mollusca	Gastropoda	Nudibranchia	Polyceridae	*Limacia*	*clavigera*	(O. F. Müller, 1776)	✓	✓
Mollusca	Gastropoda	Nudibranchia	Polyceridae	*Polycera*	*quadrilineata*	(O. F. Müller, 1776)	×	✓
Mollusca	Gastropoda	Nudibranchia	Tergipedidae	*Tergipes*	*tergipes*	(Forsskål in Niebuhr, 1775)	×	✓
Mollusca	Gastropoda	Nudibranchia	Trinchesiidae	*Trinchesia*	*caerulea*	(Montagu, 1804)	×	✓
Mollusca	Gastropoda	Nudibranchia	Trinchesiidae	*Trinchesia*	*foliata*	(Forbes & Goodsir, 1839)	×	✓
Mollusca	Gastropoda	Pleurobranchida	Pleurobranchidae	*Berthella*	*plumula*	(Montagu, 1803)	×	✓
Mollusca	Gastropoda	Trochida	Calliostomatidae	*Calliostoma*	*granulatum*	(Born, 1778)	×	✓

Mollusca	Gastropoda	Trochida	Phasianellidae	*Tricolia*	*pullus*	(Dautzenberg, 1889)	×	✓
Mollusca	Gastropoda	Trochida	Phasianellidae	*Tricolia*	*pullus*	F. Nordsieck, 1973	×	✓
Mollusca	Gastropoda	Trochida	Trochidae	*Jujubinus*	*striatus*	(Linnaeus, 1758)	×	✓
Mollusca	Gastropoda	Trochida	Trochidae	*Osilinus*	*lineatus*	(da Costa, 1778)	✓	×
Mollusca	Gastropoda	Trochida	Trochidae	*Steromphala*	*umbilicalis*	(da Costa, 1778)	✓	✓
Mollusca	Gastropoda	-	Patellidae	*Patella*	*vulgata*	Linnaeus, 1758	×	✓
Mollusca	Gastropoda	-	Plakobranchidae	*Elysia*	*viridis*	(Montagu, 1804)	×	✓
Mollusca	Polyplacophora	Chitonida	Acanthochitonidae	*Acanthochitona*	*crinita*	(Pennant, 1777)	✓	✓
Mollusca	Polyplacophora	Chitonida	Acanthochitonidae	*Acanthochitona*	*fascicularis*	(Linnaeus, 1767)	×	✓
Mollusca	Polyplacophora	Chitonida	Callochitonidae	*Callochiton*	*septemvalvis*	(Montagu, 1803)	×	✓
Mollusca	Polyplacophora	Chitonida	Lepidochitonidae	*Lepidochitona*	*cinerea*	(Linnaeus, 1767)	×	✓
Mollusca	Polyplacophora	Lepidopleurida	Leptochitonidae	*Lepidopleurus*	*cajetanus*	(Poli, 1791)	×	✓
Nemertea	Hoplonemertea	Monostilifera	Emplectonematidae	*Emplectonema*	*gracile*	(Johnston, 1837)	×	✓
Nemertea	Hoplonemertea	Monostilifera	Tetrastemmatidae	*Tetrastemma*	*candidum*	(Müller, 1774)	×	✓
Nemertea	Hoplonemertea	Monostilifera	Tetrastemmatidae	*Tetrastemma*	*coronatum*	(Quatrefages, 1846)	×	✓
Nemertea	Hoplonemertea	Monostilifera	Tetrastemmatidae	*Tetrastemma*	*flavidum*	Ehrenberg, 1828	×	✓
Nemertea	Hoplonemertea	Monostilifera	Tetrastemmatidae	*Tetrastemma*	*vermiculus*	(Quatrefages, 1846)	×	✓
Nemertea	Hoplonemertea	Monostilifera	-	*Vieitezia*	*luzmurubeae*	Junoy, Andrade & Giribet, 2010	×	✓
Nemertea	Palaeonemertea	-	Cephalotrichidae	*Cephalothrix*	*rufifrons*	(Johnston, 1837)	×	✓
Nemertea	Pilidiophora	Heteronemertea	Lineidae	*Siphonenteron*	*bilineatum*	Meneghini in Renier, 1847	×	✓
Platyhelminthes	-	Prolecithophora	Plagiostomidae	*Vorticeros*	*auriculatum*	(Müller OF, 1784)	×	✓

Table S.4.3. Taxonomic diversity (TD) and Average Taxonomic Distinctiveness (AvTD) based on the taxonomic information provided of the marine macrozoobenthic species detected through morphology in each replicate of artificial substrates (slate, PVC and granite) among sampling times. 3M – 3 months; 7M – 7 months; 10M – 10 months; 15M – 15 months.

Sampling time	Substrate	Replicate	TD	AvTD
3M	Slate	1	81,66	90,17
		2	82,06	91,73
		3	79,99	92,23
	PVC	1	83,97	90,68
		2	79,52	87,85
		3	75,71	90,33
	Granite	1	78,4	95,67
		2	79,92	90,75
		3	77,85	89,61
7M	Slate	1	86,26	89,88
		2	85,17	88,78
		3	87,86	91,27
	PVC	1	84,69	90,69
		2	87,38	93,04
		3	84,23	91,54
	Granite	1	87,98	94
		2	86,85	93,85
		3	84,46	91,87
10M	Slate	1	86,33	93,45
		2	81,86	87
		3	81,9	86,08
	PVC	1	70,05	74,99
		2	78,59	86,04

		3	70,54	76,07
	Granite	1	64,48	69,74
		2	73,58	77,76
		3	84,76	89,2
15M	Slate	1	72,29	95,93
		2	72,12	98,2
	PVC	1	77,25	97,74
		2	41	95,35
		3	73,06	91,31
	Granite	1	81,93	96,02

Table S.4.4. SIMPER analysis showing the percentage contribution of the most important phyla for differences between communities across all sampling times. 3M – 3 months; 7M – 7 months; 10M – 10 months; 15M – 15 months.

Sampling times comparison	Annelida	Bryozoa	Crustacea	Echinodermata	Hydrozoa	Mollusca	Nemertea
3M/7M	4.64	1.16	4.64	1.16	4.64	11.6	1.16
3M/10M	1.46	1.46	13.14	1.46	1.46	7.30	1.46
3M/15M	3.20	4.79	6.40	-	-	4.79	-
7M/10M	1.52	-	9.12	-	1.52	6.08	-
7M/15M	3.44	3.44	6.90	-	1.15	14.95	-
10M/15M	1.69	3.38	8.45	-	-	5.07	-

Table S.4.5. Average Taxonomic Distinctiveness (AvTD) based on the taxonomic information provided of the marine macrozoobenthic species detected in artificial substrates (slate, PVC and granite) among sampling times. 3M – 3 months; 7M – 7 months; 10M – 10 months; 15M – 15 months.

Sample	AvTD
Slate 3M	92.47
PVC 3M	94.09
Granite 3M	93.5
Slate 7M	93.12
PVC 7M	93.39
Granite 7M	93.28
Slate 10M	92.39
PVC 10M	92.52
Granite 10M	92.7
Slate 15M	91.54
PVC 15M	92.91
Granite 15M	93.34

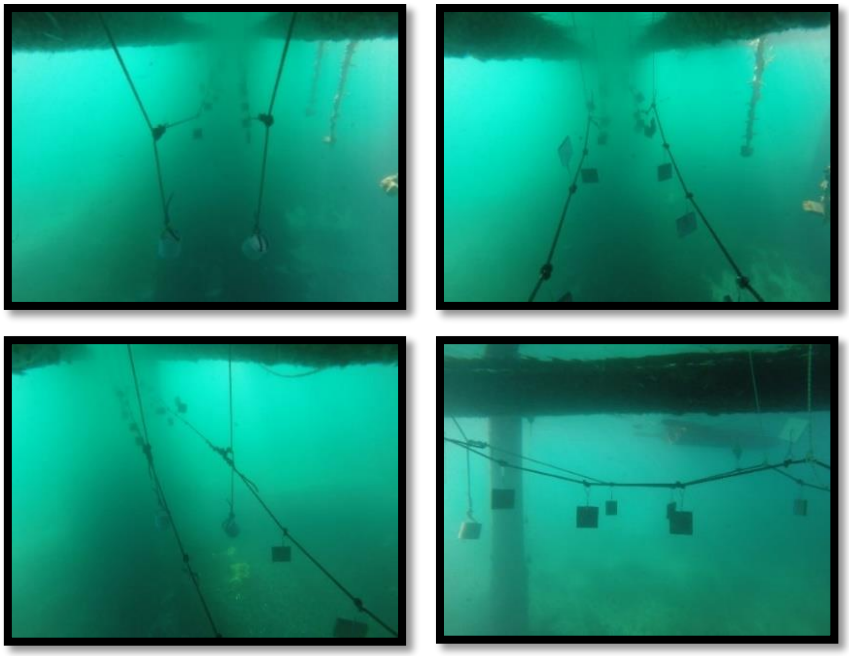


Figure S.4.1. Sampling set-up: substrates suspended horizontally and deployed in December 2016 at Toralla Island (NW Iberian Peninsula).

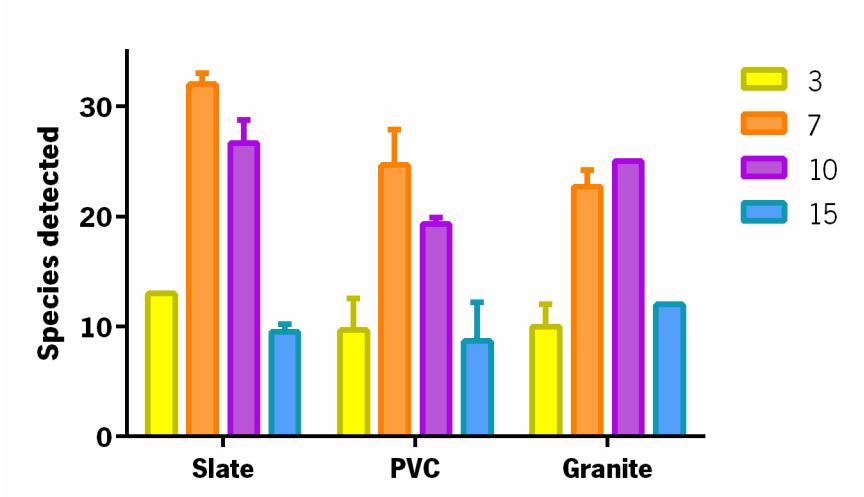


Figure S.4.2. Mean of species detected using morphology, between three replicates of artificial substrates (slate, PVC and granite) among sampling times (3, 7, 10 and 15 months).

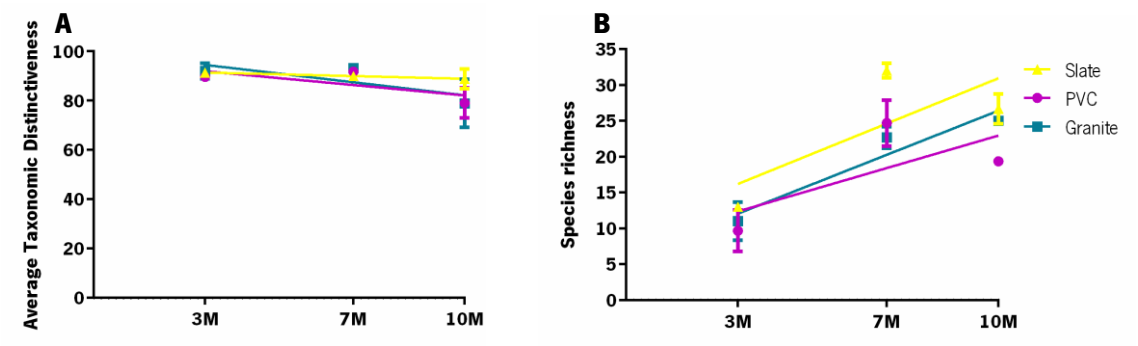


Figure S.4.3. Relation between (A) Average Taxonomic Distinctiveness and sampling times and (B) species richness and sampling times, for macrozoobenthic species identified through morphology in each replicate of artificial substrates (slate, PVC and granite). 3M – 3 months; 7 – 7 months; 10M – 10 months; 15M – 15 months. (A) Slate: r = 0.20, P = 0.22; PVC: r = 0.40, P = 0.07; Granite: r = 0.39, P = 0.07. (B) Slate: r = 0.56, P = 0.02; PVC: r = 0.44, P = 0.05; Granite: r = 0.87, P = 0.0002.

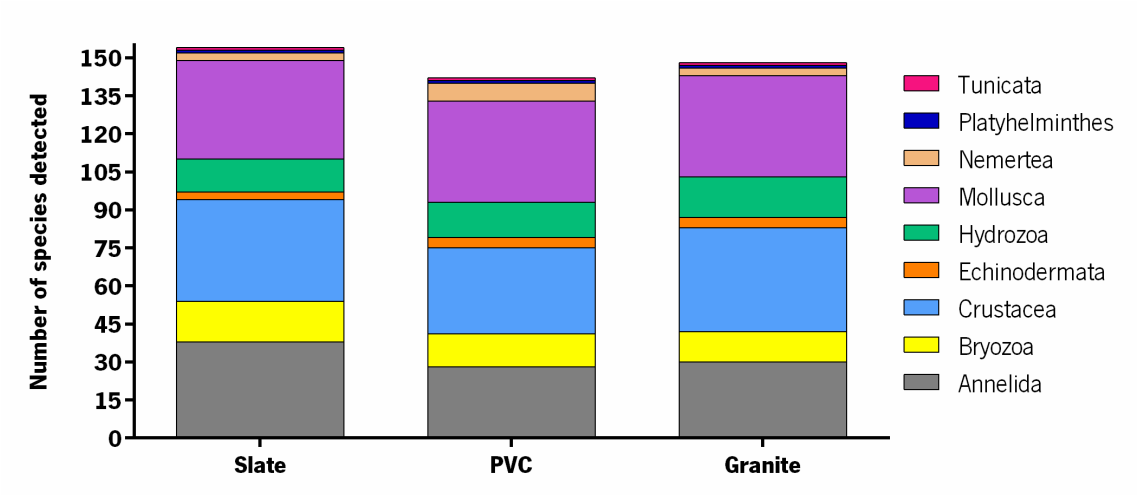


Figure S.4.4. Total number of species detected for each substrate type, pooling data from both approaches (morphology and DNA metabarcoding), and all sampling times.

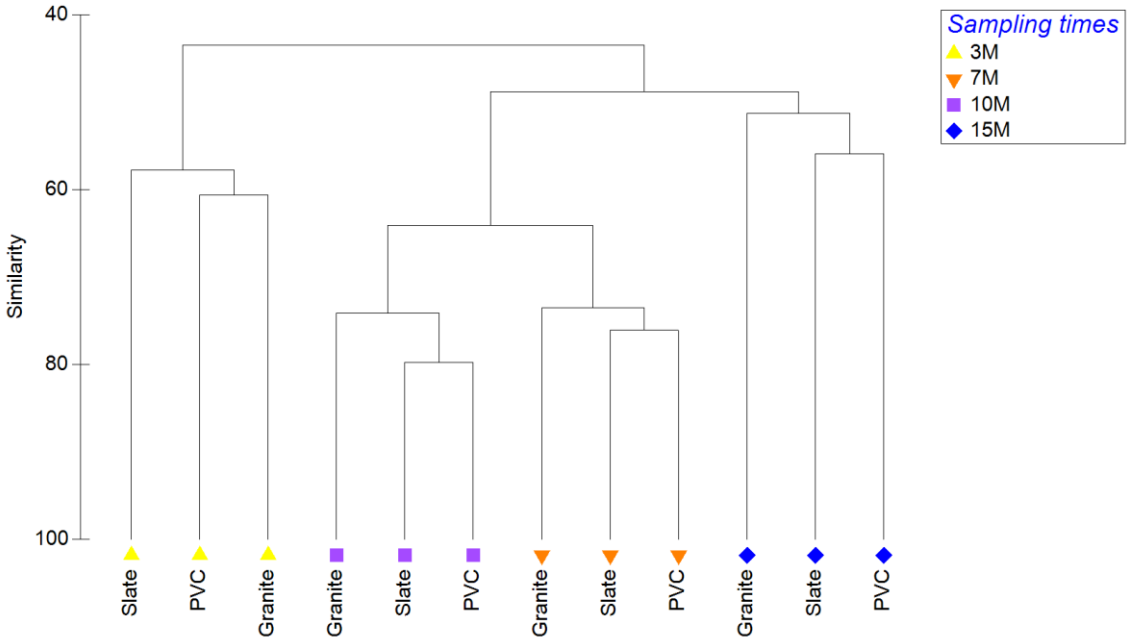


Figure S.4.5. Cluster analysis, based on Bray-Curtis similarity index, of marine macroinvertebrate species detected in the artificial substrate (Slate, PVC and Granite) among all sampling times. Sampling times: 3M – 3 months; 7M – 7 months; 10M – 10 months; 15M – 15 months.

Annexes of Chapter 5

Table S.5.1. Taxonomic classification of the marine macroinvertebrates detected through DNA metabarcoding at species-level in ARMS and ASMS substrates, with the associated authority based on WoRMS (consulted on 13rd October 2020).

Phylum	Class	Order	Family	Genus	Species	Scientific name	Authority
Annelida	Polychaeta	-	Capitellidae	*Capitella*	*teleta*	*Capitella teleta*	Blake, Grassle & Eckelbarger, 2009
Annelida	Polychaeta	-	Capitellidae	*Heteromastus*	*filiformis*	*Heteromastus filiformis*	(Claparède, 1864)
Annelida	Polychaeta	-	Chaetopteridae	*Chaetopterus*	*variopedatus*	*Chaetopterus variopedatus*	(Renier, 1804)
Annelida	Polychaeta	-	Sabellariidae	*Gunnarea*	*gaimardi*	*Gunnarea gaimardi*	(Quatrefages, 1848)
Annelida	Polychaeta	-	Sabellariidae	*Sabellaria*	*spinulosa*	*Sabellaria spinulosa*	(Leuckart, 1849)
Annelida	Polychaeta	Eunicida	Dorvilleidae	*Ophryotrocha*	*puerilis*	*Ophryotrocha puerilis*	Claparède & Mecznikow, 1869
Annelida	Polychaeta	Eunicida	Lumbrineridae	*Ninoe*	*nigripes*	*Ninoe nigripes*	Verrill, 1873
Annelida	Polychaeta	Phyllodocida	Chrysopetalidae	*Bhawania*	*cryptocephala*	*Bhawania cryptocephala*	Gravier, 1901
Annelida	Polychaeta	Phyllodocida	Hesionidae	*Psamathe*	*fusca*	*Psamathe fusca*	Johnston, 1836
Annelida	Polychaeta	Phyllodocida	Hesionidae	*Syllidia*	*armata*	*Syllidia armata*	Quatrefages, 1866
Annelida	Polychaeta	Phyllodocida	Nereididae	*Laeonereis*	*acuta*	*Laeonereis acuta*	(Treadwell, 1923)
Annelida	Polychaeta	Phyllodocida	Nereididae	*Perinereis*	*aibuhitensis*	*Perinereis aibuhitensis*	(Grube, 1878)

Annelida	Polychaeta	Phyllodocida	Nereididae	*Platynereis*	*dumerilii*	*Platynereis dumerilii*	(Audouin & Milne Edwards, 1833)
Annelida	Polychaeta	Phyllodocida	Phyllodocidae	*Eumida*	*arctica*	*Eumida arctica*	(Annenkova, 1946)
Annelida	Polychaeta	Phyllodocida	Phyllodocidae	*Eumida*	*sanguinea*	*Eumida sanguinea*	(Örsted, 1843)
Annelida	Polychaeta	Phyllodocida	Phyllodocidae	*Pterocirrus*	*macroceros*	*Pterocirrus macroceros*	(Grube, 1860)
Annelida	Polychaeta	Phyllodocida	Polynoidae	*Alentia*	*gelatinosa*	*Alentia gelatinosa*	(M. Sars, 1835)
Annelida	Polychaeta	Phyllodocida	Polynoidae	*Harmothoe*	*impar*	*Harmothoe impar*	(Johnston, 1839)
Annelida	Polychaeta	Phyllodocida	Polynoidae	*Lepidonotus*	*clava*	*Lepidonotus clava*	(Montagu, 1808)
Annelida	Polychaeta	Phyllodocida	Sigalionidae	*Sthenelais*	*boa*	*Sthenelais boa*	(Johnston, 1833)
Annelida	Polychaeta	Phyllodocida	Syllidae	*Eusyllis*	*blomstrandi*	*Eusyllis blomstrandi*	Malmgren, 1867
Annelida	Polychaeta	Phyllodocida	Syllidae	*Exogone*	*naidinoides*	*Exogone naidinoides*	Westheide, 1974
Annelida	Polychaeta	Phyllodocida	Syllidae	*Myrianida*	*brachycephala*	*Myrianida brachycephala*	(Marenzeller, 1874)
Annelida	Polychaeta	Phyllodocida	Syllidae	*Myrianida*	*edwarsi*	*Myrianida edwarsi*	(Saint Joseph, 1887)
Annelida	Polychaeta	Phyllodocida	Syllidae	*Myrianida*	*inermis*	*Myrianida inermis*	(Saint Joseph, 1887)
Annelida	Polychaeta	Phyllodocida	Syllidae	*Myrianida*	*pinnigera*	*Myrianida pinnigera*	(Montagu, 1808)
Annelida	Polychaeta	Phyllodocida	Syllidae	*Myrianida*	*prolifera*	*Myrianida prolifera*	(O.F. Müller, 1788)
Annelida	Polychaeta	Phyllodocida	Syllidae	*Nudisyllis*	*pulligera*	*Nudisyllis pulligera*	(Krohn, 1852)
Annelida	Polychaeta	Phyllodocida	Syllidae	*Prosphaerosyllis*	*longipapillata*	*Prosphaerosyllis longipapillata*	(Hartmann-Schröder, 1979)

Annelida	Polychaeta	Phyllodocida	Syllidae	*Sphaerosyllis*	*pirifera*	*Sphaerosyllis pirifera*	Claparède, 1868
Annelida	Polychaeta	Sabellida	Sabellidae	*Sabella*	*spallanzanii*	*Sabella spallanzanii*	(Gmelin, 1791)
Annelida	Polychaeta	Sabellida	Serpulidae	*Hydroides*	*ezoensis*	*Hydroides ezoensis*	Okuda, 1934
Annelida	Polychaeta	Spionida	Spionidae	*Aonides*	*oxycephala*	*Aonides oxycephala*	(Sars, 1862)
Annelida	Polychaeta	Spionida	Spionidae	*Laonice*	*cirrata*	*Laonice cirrata*	(M. Sars, 1851)
Annelida	Polychaeta	Spionida	Spionidae	*Prionospio*	*fallax*	*Prionospio fallax*	Söderström, 1920
Annelida	Polychaeta	Terebellida	Cirratulidae	*Cirratulus*	*spectabilis*	*Cirratulus spectabilis*	(Kinberg, 1866)
Annelida	Polychaeta	Terebellida	Terebellidae	*Polycirrus*	*carolinensis*	*Polycirrus carolinensis*	Day, 1973
Annelida	Polychaeta	Terebellida	Terebellidae	*Terebella*	*lapidaria*	*Terebella lapidaria*	Linnaeus, 1767
Arthropoda	Hexanauplia	Calanoida	Calanidae	*Calanus*	*glacialis*	*Calanus glacialis*	Jaschnov, 1955
Arthropoda	Hexanauplia	Calanoida	Centropagidae	*Centropages*	*typicus*	*Centropages typicus*	Krøyer, 1849
Arthropoda	Hexanauplia	Calanoida	Centropagidae	*Isias*	*clavipes*	*Isias clavipes*	Boeck, 1865
Arthropoda	Hexanauplia	Calanoida	Paracalanidae	*Paracalanus*	*parvus*	*Paracalanus parvus*	(Claus, 1863)
Arthropoda	Hexanauplia	Calanoida	Temoridae	*Temora*	*longicornis*	*Temora longicornis*	(Müller O.F., 1785)
Arthropoda	Hexanauplia	Cyclopoida	Cyclopinidae	*Cyclopina*	*gracilis*	*Cyclopina gracilis*	Claus, 1863
Arthropoda	Hexanauplia	Harpacticoida	Dactylopusiidae	*Dactylopusia*	*pauciarticulata*	*Dactylopusia pauciarticulata*	Chang & Song, 1997
Arthropoda	Hexanauplia	Harpacticoida	Miraciidae	*Paramphiascella*	*fulvofasciata*	*Paramphiascella fulvofasciata*	Rosenfield & Coull, 1974

Arthropoda	Hexanauplia	Harpacticoida	Miraciidae	*Typhlamphiascus*	*typhlops*	*Typhlamphiascus typhlops*	(Sars G.O., 1906)
Arthropoda	Hexanauplia	Sessilia	Balanidae	*Balanus*	*trigonus*	*Balanus trigonus*	Darwin, 1854
Arthropoda	Hexanauplia	Sessilia	Balanidae	*Perforatus*	*perforatus*	*Perforatus perforatus*	(Bruguière, 1789)
Arthropoda	Hexanauplia	Sessilia	Verrucidae	*Verruca*	*stroemia*	*Verruca stroemia*	(O.F. Müller, 1776)
Arthropoda	Malacostraca	Amphipoda	Ampithoidae	*Ampithoe*	*rubricata*	*Ampithoe rubricata*	(Montagu, 1808)
Arthropoda	Malacostraca	Amphipoda	Aoridae	*Aora*	*gracilis*	*Aora gracilis*	(Spence Bate, 1857)
Arthropoda	Malacostraca	Amphipoda	Aoridae	*Aora*	*typica*	*Aora typica*	Krøyer, 1845
Arthropoda	Malacostraca	Amphipoda	Aoridae	*Microdeutopus*	*chelifer*	*Microdeutopus chelifer*	(Spence Bate, 1862)
Arthropoda	Malacostraca	Amphipoda	Caprellidae	*Caprella*	*acanthifera*	*Caprella acanthifera*	Leach, 1814
Arthropoda	Malacostraca	Amphipoda	Caprellidae	*Phtisica*	*marina*	*Phtisica marina*	Slabber, 1769
Arthropoda	Malacostraca	Amphipoda	Corophiidae	*Corophium*	*multisetosum*	*Corophium multisetosum*	Stock, 1952
Arthropoda	Malacostraca	Amphipoda	Corophiidae	*Monocorophium*	*sextonae*	*Monocorophium sextonae*	(Crawford, 1937)
Arthropoda	Malacostraca	Amphipoda	Ischyroceridae	*Ericthonius*	*punctatus*	*Ericthonius punctatus*	(Spence Bate, 1857)
Arthropoda	Malacostraca	Amphipoda	Ischyroceridae	*Jassa*	*herdmani*	*Jassa herdmani*	(Walker, 1893)
Arthropoda	Malacostraca	Amphipoda	Ischyroceridae	*Jassa*	*slatteryi*	*Jassa slatteryi*	Conlan, 1990
Arthropoda	Malacostraca	Decapoda	Calappidae	*Calappa*	*bilineata*	*Calappa bilineata*	P.K.L. Ng, J.C.Y. Lai & Aungtonya, 2002
Arthropoda	Malacostraca	Decapoda	Crangonidae	*Crangon*	*crangon*	*Crangon crangon*	(Linnaeus, 1758)

Arthropoda	Malacostraca	Decapoda	Hippolytidae	*Hippolyte*	*varians*	*Hippolyte varians*	Leach, 1814 [in Leach, 1813-1815]
Arthropoda	Malacostraca	Decapoda	Inachidae	*Inachus*	*phalangium*	*Inachus phalangium*	(Fabricius, 1775)
Arthropoda	Malacostraca	Decapoda	Inachidae	*Macropodia*	*parva*	*Macropodia rostrata*	(Linnaeus, 1761)
Arthropoda	Malacostraca	Decapoda	Oregoniidae	*Hyas*	*araneus*	*Hyas araneus*	(Linnaeus, 1758)
Arthropoda	Malacostraca	Decapoda	Paguridae	*Pagurus*	*cuanensis*	*Pagurus cuanensis*	W. Thompson, 1844
Arthropoda	Malacostraca	Decapoda	Pilumnidae	*Pilumnus*	*hirtellus*	*Pilumnus hirtellus*	(Linnaeus, 1761)
Arthropoda	Malacostraca	Decapoda	Polybiidae	*Liocarcinus*	*navigator*	*Liocarcinus navigator*	(Herbst, 1794)
Arthropoda	Malacostraca	Decapoda	Polybiidae	*Necora*	*puber*	*Necora puber*	(Linnaeus, 1767)
Arthropoda	Malacostraca	Decapoda	Porcellanidae	*Pisidia*	*longicornis*	*Pisidia longicornis*	(Linnaeus, 1767)
Arthropoda	Malacostraca	Decapoda	Thoridae	*Eualus*	*cranchii*	*Eualus cranchii*	(Leach, 1817 [in Leach, 1815-1875])
Arthropoda	Malacostraca	Decapoda	Thoridae	*Eualus*	*occultus*	*Eualus occultus*	(Lebour, 1936)
Arthropoda	Malacostraca	Decapoda	Xanthidae	*Xantho*	*hydrophilus*	*Xantho hydrophilus*	(Herbst, 1790)
Arthropoda	Malacostraca	Isopoda	Arcturidae	*Astacilla*	*damnoniensis*	*Astacilla damnoniensis*	(Stebbing, 1874)
Arthropoda	Malacostraca	Isopoda	Hologнathidae	*Cleantis*	*prismatica*	*Cleantis prismatica*	(Risso, 1826)
Arthropoda	Malacostraca	Mysida	Mysidae	*Siriella*	*jaltensis*	*Siriella jaltensis*	Czerniavsky, 1868
Arthropoda	Ostracoda	Myodocopida	Cylindroleberididae	*Parasterope*	*pollex*	*Parasterope pollex*	Kornicker in Bowman & Kornicker, 1967
Arthropoda	Ostracoda	Podocopida	Hemicytheridae	*Coquimba*	*ishizakii*	*Coquimba ishizakii*	Yajima, 1978

Arthropoda	Ostracoda	Podocopida	Loxoconchidae	*Loxocorniculum*	*mutsuense*	*Loxocorniculum mutsuense*	Ishizaki, 1971
Arthropoda	Ostracoda	Podocopida	Pontocyprididae	*Ekpontocypris*	*pirifera*	*Ekpontocypris pirifera*	(Mueller, 1894)
Arthropoda	Ostracoda	Podocopida	Trachyleberididae	*Pistocythereis*	*bradyformis*	*Pistocythereis bradyformis*	(Ishizaki, 1968) Gou in Gou, Zheng & Huang, 1983
Arthropoda	Pycnogonida	Pantopoda	Ammotheidae	*Achelia*	*echinata*	*Achelia echinata*	Hodge, 1864
Bryozoa	Gymnolaemata	Cheilostomatida	Candidae	*Scrupocellaria*	*scruposa*	*Scrupocellaria scruposa*	(Linnaeus, 1758)
Bryozoa	Gymnolaemata	Cheilostomatida	Candidae	*Tricellaria*	*inopinata*	*Tricellaria inopinata*	d'Hondt & Occhipinti Ambrogi, 1985
Bryozoa	Gymnolaemata	Cheilostomatida	Celleporidae	*Celleporina*	*caliciformis*	*Celleporina caliciformis*	(Lamouroux, 1816)
Bryozoa	Gymnolaemata	Cheilostomatida	Celleporidae	*Galeopsis*	*porcellanicus*	*Galeopsis porcellanicus*	(Hutton, 1873)
Bryozoa	Gymnolaemata	Cheilostomatida	Cryptosulidae	*Cryptosula*	*pallasiana*	*Cryptosula pallasiana*	(Moll, 1803)
Bryozoa	Gymnolaemata	Cheilostomatida	Electridae	*Electra*	*pilosa*	*Electra pilosa*	(Linnaeus, 1767)
Bryozoa	Gymnolaemata	Cheilostomatida	Escharinidae	*Chiastosella*	*watersi*	*Chiastosella watersi*	Stach, 1937
Bryozoa	Gymnolaemata	Cheilostomatida	Exochellidae	*Escharoides*	*coccinea*	*Escharoides coccinea*	(Abildgaard, 1806)
Bryozoa	Gymnolaemata	Cheilostomatida	Haplopomidae	*Haplopoma*	*graniferum*	*Haplopoma graniferum*	(Johnston, 1847)
Bryozoa	Gymnolaemata	Cheilostomatida	Microporellidae	*Microporella*	*ciliata*	*Microporella ciliata*	(Pallas, 1766)
Bryozoa	Gymnolaemata	Cheilostomatida	Microporellidae	*Microporella*	*ordo*	*Microporella ordo*	Brown, 1952
Bryozoa	Gymnolaemata	Cheilostomatida	Scruparidae	*Scruparia*	*chelata*	*Scruparia chelata*	(Linnaeus, 1758)
Bryozoa	Gymnolaemata	Cheilostomatida	Watersiporidae	*Watersipora*	*subtorquata*	*Watersipora subtorquata*	(d'Orbigny, 1852)

Bryozoa	Stenolaemata	Cyclostomatida	Crisiidae	*Crisia*	*aculeata*	*Crisia aculeata*	Hassall, 1841
Bryozoa	Stenolaemata	Cyclostomatida	Tubuliporidae	*Tubulipora*	*liliacea*	*Tubulipora liliacea*	(Pallas, 1766)
Bryozoa	Stenolaemata	Cyclostomatida	Tubuliporidae	*Tubulipora*	*lobifera*	*Tubulipora lobifera*	Hastings, 1963
Chordata	Ascidiacea	Aplousobranchia	Clavelinidae	*Pycnoclavella*	*detorta*	*Pycnoclavella detorta*	(Sluiter, 1904)
Chordata	Ascidiacea	Phlebobranchia	Cionidae	*Ciona*	*intestinalis*	*Ciona intestinalis*	(Linnaeus, 1767)
Chordata	Ascidiacea	Phlebobranchia	Cionidae	*Ciona*	*savignyi*	*Ciona savignyi*	Herdman, 1882
Chordata	Ascidiacea	Stolidobranchia	Molgulidae	*Molgula*	*complanata*	*Molgula complanata*	Alder & Hancock, 1870
Chordata	Ascidiacea	Stolidobranchia	Styelidae	*Asterocarpa*	*humilis*	*Asterocarpa humilis*	(Heller, 1878)
Chordata	Ascidiacea	Stolidobranchia	Styelidae	*Botryllus*	*schlosseri*	*Botryllus schlosseri*	(Pallas, 1766)
Chordata	Ascidiacea	Stolidobranchia	Styelidae	*Metandrocarpa*	*taylori*	*Metandrocarpa taylori*	Huntsman, 1912
Chordata	Ascidiacea	Stolidobranchia	Styelidae	*Polycarpa*	*pomaria*	*Polycarpa pomaria*	(Savigny, 1816)
Chordata	Ascidiacea	Stolidobranchia	Styelidae	*Symplegma*	*viride*	*Symplegma viride*	Herdman, 1886
Cnidaria	Hydrozoa	Anthoathecata	Bougainvilliidae	*Bougainvillia*	*muscus*	*Bougainvillia muscus*	(Allman, 1863)
Cnidaria	Hydrozoa	Anthoathecata	Bougainvilliidae	*Garveia*	*grisea*	*Garveia grisea*	(Motz-Kossowska, 1905)
Cnidaria	Hydrozoa	Anthoathecata	Cladonematidae	*Cladonema*	*californicum*	*Cladonema californicum*	Hyman, 1947
Cnidaria	Hydrozoa	Anthoathecata	Corynidae	*Coryne*	*pusilla*	*Coryne pusilla*	Gaertner, 1774
Cnidaria	Hydrozoa	Anthoathecata	Eudendriidae	*Eudendrium*	*racemosum*	*Eudendrium racemosum*	(Cavolini, 1785)
Cnidaria	Hydrozoa	Anthoathecata	Hydractiniidae	*Podocoryna*	*exigua*	*Podocoryna exigua*	(Haeckel, 1880)

Cnidaria	Hydrozoa	Anthoathecata	Pandeidae	*Amphinema*	*dinema*	*Amphinema dinema*	(Péron & Lesueur, 1810)
Cnidaria	Hydrozoa	Anthoathecata	Tubulariidae	*Ectopleura*	*marina*	*Ectopleura marina*	(Torrey, 1902)
Cnidaria	Hydrozoa	Leptothecata	Aglaopheniidae	*Cladocarpus*	*integer*	*Cladocarpus integer*	(Sars, 1873)
Cnidaria	Hydrozoa	Leptothecata	Campanulariidae	*Campanularia*	*hincksii*	*Campanularia hincksii*	Alder, 1856
Cnidaria	Hydrozoa	Leptothecata	Campanulariidae	*Clytia*	*gracilis*	*Clytia gracilis*	(Sars, 1850)
Cnidaria	Hydrozoa	Leptothecata	Campanulariidae	*Clytia*	*hemisphaerica*	*Clytia hemisphaerica*	(Linnaeus, 1767)
Cnidaria	Hydrozoa	Leptothecata	Campanulariidae	*Clytia*	*paulensis*	*Clytia paulensis*	(Vanhöffen, 1910)
Cnidaria	Hydrozoa	Leptothecata	Campanulariidae	*Obelia*	*dichotoma*	*Obelia dichotoma*	(Linnaeus, 1758)
Cnidaria	Hydrozoa	Leptothecata	Campanulariidae	*Obelia*	*geniculata*	*Obelia geniculata*	(Linnaeus, 1758)
Cnidaria	Hydrozoa	Leptothecata	Campanulariidae	*Orthopyxis*	*everta*	*Orthopyxis everta*	(Clark, 1876)
Cnidaria	Hydrozoa	Leptothecata	Clathrozoidae	*Clathrozoon*	*wilsoni*	*Clathrozoon wilsoni*	Spencer, 1891
Cnidaria	Hydrozoa	Leptothecata	Haleciidae	*Halecium*	*beanii*	*Halecium beanii*	(Johnston, 1838)
Cnidaria	Hydrozoa	Leptothecata	Haleciidae	*Halecium*	*halecinum*	*Halecium halecinum*	(Linnaeus, 1758)
Cnidaria	Hydrozoa	Leptothecata	Haleciidae	*Halecium*	*mediterraneum*	*Halecium mediterraneum*	Weismann, 1883
Cnidaria	Hydrozoa	Leptothecata	Haleciidae	*Halecium*	*pusillum*	*Halecium pusillum*	Sars, 1856
Cnidaria	Hydrozoa	Leptothecata	Halopteridae	*Halopteris*	*catharina*	*Halopteris catharina*	(Johnston, 1833)

Cnidaria	Hydrozoa	Leptothecata	Kirchenpaueriidae	*Kirchenpaueria*	*pinnata*	*Kirchenpaueria pinnata*	(Linnaeus, 1758)
Cnidaria	Hydrozoa	Leptothecata	Laodiceidae	*Laodicea*	*undulata*	*Laodicea undulata*	(Forbes & Goodsir, 1853)
Cnidaria	Hydrozoa	Leptothecata	Melicertidae	*Melicertum*	*octocostatum*	*Melicertum octocostatum*	(M. Sars, 1835)
Cnidaria	Hydrozoa	Leptothecata	Mitrocomidae	*Mitrocomella*	*brownei*	*Mitrocomella brownei*	(Kramp, 1930)
Cnidaria	Hydrozoa	Leptothecata	Mitrocomidae	*Mitrocomella*	*niwai*	*Mitrocomella niwai*	Bouillon & Barnett, 1999
Cnidaria	Hydrozoa	Leptothecata	Phialellidae	*Phialella*	*quadrata*	*Phialella quadrata*	(Forbes, 1848)
Cnidaria	Hydrozoa	Leptothecata	Phylactothecidae	*Hydrodendron*	*mirabile*	*Hydrodendron mirabile*	(Hincks, 1866)
Cnidaria	Hydrozoa	Leptothecata	Plumulariidae	*Nemertesia*	*antennina*	*Nemertesia antennina*	(Linnaeus, 1758)
Cnidaria	Hydrozoa	Leptothecata	Sertularellidae	*Sertularella*	*ellisii*	*Sertularella ellisii*	(Deshayes & Milne Edwards, 1836)
Cnidaria	Hydrozoa	Leptothecata	Sertularellidae	*Sertularella*	*polyzonias*	*Sertularella polyzonias*	(Linnaeus, 1758)
Cnidaria	Hydrozoa	Leptothecata	Sertulariidae	*Abietinaria*	*filicula*	*Abietinaria filicula*	(Ellis & Solander, 1786)
Cnidaria	Hydrozoa	Leptothecata	Sertulariidae	*Amphisbetia*	*operculata*	*Amphisbetia operculata*	(Linnaeus, 1758)
Echinodermata	Asteroidea	Forcipulatida	Asteriidae	*Asterias*	*rubens*	*Asterias rubens*	Linnaeus, 1758
Echinodermata	Asteroidea	Forcipulatida	Asteriidae	*Diplasterias*	*brucei*	*Diplasterias brucei*	(Koehler, 1907)
Echinodermata	Asteroidea	Forcipulatida	Asteriidae	*Marthasterias*	*glacialis*	*Marthasterias glacialis*	(Linnaeus, 1758)

Echinodermata	Asteroidea	Paxillosida	Astropectinidae	*Astropecten*	*articulatus*	*Astropecten articulatus*	(Say, 1825)
Echinodermata	Asteroidea	Valvatida	Acanthasteridae	*Acanthaster*	*planci*	*Acanthaster planci*	(Linnaeus, 1758)
Echinodermata	Asteroidea	Valvatida	Archasteridae	*Archaster*	*typicus*	*Archaster typicus*	Müller & Troschel, 1840
Echinodermata	Asteroidea	Valvatida	Asterinidae	*Asterina*	*gibbosa*	*Asterina gibbosa*	(Pennant, 1777)
Echinodermata	Asteroidea	Valvatida	Asterinidae	*Patiria*	*pectinifera*	*Patiria pectinifera*	(Muller & Troschel, 1842)
Echinodermata	Crinoidea	Comatulida	Antedonidae	*Antedon*	*bifida*	*Antedon bifida*	(Pennant, 1777)
Echinodermata	Crinoidea	Comatulida	Colobometridae	*Cyllometra*	*manca*	*Cyllometra manca*	(Carpenter, 1888)
Echinodermata	Crinoidea	Isocrinida	Isselicrinidae	*Metacrinus*	*levii*	*Metacrinus levii*	Améziane-Cominardi, 1990
Echinodermata	Echinoidea	Camarodonta	Parechinidae	*Psammechinus*	*miliaris*	*Psammechinus miliaris*	(P.L.S. Müller, 1771)
Echinodermata	Echinoidea	Camarodonta	Strongylocentrotidae	*Strongylocentrotus*	*purpuratus*	*Strongylocentrotus purpuratus*	(Stimpson, 1857)

Echinodermata	Echinoidea	Spatangoida	Schizasteridae	*Abatus*	*cavernosus*	*Abatus cavernosus*	(Philippi, 1845)
Echinodermata	Holothuroidea	Dendrochirotida	Phyllophoridae	*Thyone*	*fusus*	*Thyone fusus*	(O.F. Müller, 1776)
Echinodermata	Ophiuroidea	Amphilepidida	Amphiuridae	*Amphipholis*	*squamata*	*Amphipholis squamata*	(Delle Chiaje, 1828)
Echinodermata	Ophiuroidea	Amphilepidida	Amphiuridae	*Amphiura*	*filiformis*	*Amphiura filiformis*	(O.F. Müller, 1776)
Echinodermata	Ophiuroidea	Amphilepidida	Ophiotrichidae	*Ophiothrix*	*fragilis*	*Ophiothrix fragilis*	(Abildgaard in O.F. Müller, 1789)
Echinodermata	Ophiuroidea	Amphilepidida	Ophiotrichidae	*Ophiothrix*	*oerstedii*	*Ophiothrix (Ophiothrix) oerstedii*	Lütken, 1856
Echinodermata	Ophiuroidea	Ophiacanthida	Ophiodermatidae	*Ophioderma*	*cinereum*	*Ophioderma cinereum*	Müller & Troschel, 1842
Echinodermata	Ophiuroidea	Ophiacanthida	Ophiotomidae	*Ophiocomina*	*nigra*	*Ophiocomina nigra*	(Abildgaard in O.F. Müller, 1789)
Mollusca	Bivalvia	Adapedonta	Hiatellidae	*Hiatella*	*arctica*	*Hiatella arctica*	(Linnaeus, 1767)
Mollusca	Bivalvia	Adapedonta	Pharidae	*Ensis*	*ensis*	*Ensis ensis*	(Linnaeus, 1758)
Mollusca	Bivalvia	Arcida	Arcidae	*Anadara*	*broughtonii*	*Anadara broughtonii*	(Schrenck, 1867)
Mollusca	Bivalvia	Mytilida	Mytilidae	*Musculus*	*lateralis*	*Musculus lateralis*	(Say, 1822)

Mollusca	Bivalvia	Mytilida	Mytilidae	*Musculus*	*subpictus*	*Musculus subpictus*	(Cantraine, 1835)
Mollusca	Bivalvia	Ostreida	Pinnidae	*Atrina*	*pectinata*	*Atrina pectinata*	(Linnaeus, 1767)
Mollusca	Bivalvia	Pectinida	Plicatulidae	*Plicatula*	*australis*	*Plicatula australis*	Lamarck, 1819
Mollusca	Gastropoda	Aplysiida	Aplysiidae	*Aplysia*	*depilans*	*Aplysia depilans*	Gmelin, 1791
Mollusca	Gastropoda	Lepetellida	Haliotidae	*Haliotis*	*midae*	*Haliotis midae*	Linnaeus, 1758
Mollusca	Gastropoda	Lepetellida	Haliotidae	*Haliotis*	*tuberculata*	*Haliotis tuberculata*	Linnaeus, 1758
Mollusca	Gastropoda	Littorinimorpha	Calyptraeidae	*Crepidula*	*navicella*	*Crepidula navicella*	(Lesson, 1831)
Mollusca	Gastropoda	Littorinimorpha	Calyptraeidae	*Crepipatella*	*dilatata*	*Crepipatella dilatata*	(Lamarck, 1822)
Mollusca	Gastropoda	Littorinimorpha	Rissoidae	*Rissoa*	*parva*	*Rissoa parva*	(da Costa, 1778)
Mollusca	Gastropoda	Neogastropoda	Fasciolariidae	*Fusolatirus*	*rikae*	*Fusolatirus rikae*	(Fraussen, 2003)
Mollusca	Gastropoda	Neogastropoda	Nassariidae	*Tritia*	*conspersa*	*Tritia conspersa*	(Philippi, 1849)
Mollusca	Gastropoda	Neogastropoda	Nassariidae	*Tritia*	*incrassata*	*Tritia incrassata*	(Strøm, 1768)
Mollusca	Gastropoda	Neogastropoda	Nassariidae	*Tritia*	*obsoleta*	*Tritia obsoleta*	(Say, 1822)
Mollusca	Gastropoda	Nudibranchia	Dotidae	*Doto*	*coronata*	*Doto coronata*	(Gmelin, 1791)
Mollusca	Gastropoda	Nudibranchia	Eubranchidae	*Amphorina*	*linensis*	*Amphorina linensis*	(Garcia-Gomez, Cervera & Garcia, 1990)
Mollusca	Gastropoda	Nudibranchia	Facelinidae	*Facelina*	*bostoniensis*	*Facelina bostoniensis*	(Couthouy, 1838)
Mollusca	Gastropoda	Nudibranchia	Onchidorididae	*Onchidoris*	*bilamellata*	*Onchidoris bilamellata*	(Linnaeus, 1767)
Mollusca	Gastropoda	Nudibranchia	Tergipedidae	*Tergipes*	*tergipes*	*Tergipes tergipes*	(Forsskål in Niebuhr, 1775)

Mollusca	Gastropoda	Nudibranchia	Trinchesiidae	*Trinchesia*	*foliata*	*Trinchesia foliata*	(Forbes & Goodsir, 1839)
Mollusca	Gastropoda	Trochida	Calliostomatidae	*Calliostoma*	*zizyphinum*	*Calliostoma zizyphinum*	(Linnaeus, 1758)
Mollusca	Gastropoda	Trochida	Phasianellidae	*Tricolia*	*pullus*	*Tricolia pullus*	(Linnaeus, 1758)
Mollusca	Gastropoda	Trochida	Trochidae	*Steromphala*	*cineraria*	*Steromphala cineraria*	(Linnaeus, 1758)
Mollusca	Polyplacophora	Chitonida	Acanthochitonidae	*Acanthochitona*	*fascicularis*	*Acanthochitona fascicularis*	(Linnaeus, 1767)
Mollusca	Polyplacophora	Chitonida	Callochitonidae	*Callochiton*	*bouveti*	*Callochiton bouveti*	Thiele, 1906
Mollusca	Polyplacophora	Chitonida	Mopaliidae	*Plaxiphora*	*albida*	*Plaxiphora albida*	(Blainville, 1825)
Nematoda	Chromadorea	Monhysterida	Monhysteridae	*Halomonhystera*	*disjuncta*	*Halomonhystera disjuncta*	(Bastian, 1865) Andrassy, 2006
Nemertea	Hoplonemertea	Monostilifera	-	*Abyssonemertes*	*kajiharai*	*Abyssonemertes kajiharai*	Chernyshev & Polyakova, 2017
Nemertea	Hoplonemertea	Monostilifera	-	*Vieitezia*	*luzmurubeae*	*Vieitezia luzmurubeae*	Junoy, Andrade & Giribet, 2010
Nemertea	Hoplonemertea	Monostilifera	Oerstediidae	*Oerstedia*	*dorsalis*	*Oerstedia dorsalis*	(Abildgaard, 1806)
Nemertea	Hoplonemertea	Monostilifera	Ototyphlonemertidae	*Ototyphlonemertes*	*macintoshi*	*Ototyphlonemertes (Macintoshi) macintoshi*	Bürger, 1895
Nemertea	Hoplonemertea	Monostilifera	Tetrastemmatidae	*Tetrastemma*	*coronatum*	*Tetrastemma coronatum*	(Quatrefages, 1846)
Nemertea	Hoplonemertea	Monostilifera	Tetrastemmatidae	*Tetrastemma*	*flavidum*	*Tetrastemma flavidum*	Ehrenberg, 1828
Nemertea	Hoplonemertea	Monostilifera	Tetrastemmatidae	*Tetrastemma*	*peltatum*	*Tetrastemma peltatum*	Bürger, 1895
Nemertea	Palaeonemertea	Archinemertea	Cephalotrichidae	*Cephalothrix*	*rufifrons*	*Cephalothrix rufifrons*	(Johnston, 1837)
Nemertea	Palaeonemertea	Tubulaniformes	Tubulanidae	*Tubulanus*	*annulatus*	*Tubulanus annulatus*	(Montagu, 1804)

Nemertea	Pilidiophora	Heteronemertea	Lineidae	*Siphonenteron*	*bilineatus*	*Siphonenteron bilineatum*	Meneghini in Renier, 1847
Platyhelminthes	-	Dolichomicrostomida	Microstomidae	*Microstomum*	*papillosum*	*Microstomum papillosum*	Graff, 1882
Platyhelminthes	-	Polycladida	Euryleptidae	*Prostheceraeus*	*vittatus*	*Prostheceraeus vittatus*	(Montagu, 1815)
Platyhelminthes	-	Polycladida	Notoplanidae	*Notoplana*	*australis*	*Notoplana australis*	(Schmarda, 1859)
Platyhelminthes	-	Rhabdocoela	Polycystididae	*Progyrator*	*mamertinus*	*Progyrator mamertinus*	(Graff, 1874)
Platyhelminthes	-	Rhabdocoela	Trigonostomidae	*Trigonostomum*	*venenosum*	*Trigonostomum venenosum*	(Uljanin, 1870)
Porifera	Demospongiae	Clionaida	Clionaidae	*Pione*	*vastifica*	*Pione vastifica*	(Hancock, 1849)
Porifera	Demospongiae	Dictyoceratida	Dysideidae	*Dysidea*	*etheria*	*Dysidea etheria*	Laubenfels, 1936
Porifera	Demospongiae	Haplosclerida	Niphatidae	*Niphates*	*erecta*	*Niphates erecta*	Duchassaing & Michelotti, 1864
Porifera	Demospongiae	Poecilosclerida	Crellidae	*Crella*	*incrustans*	*Crella incrustans*	(Carter, 1885)
Porifera	Demospongiae	Poecilosclerida	Hymedesmiidae	*Hymedesmia*	*methanophila*	*Hymedesmia (Stylopus) methanophila*	Cárdenas, 2019
Porifera	Demospongiae	Poecilosclerida	Mycalidae	*Mycale*	*sanguinea*	*Mycale (Carmia) sanguinea*	Tsurnamal, 1969
Porifera	Demospongiae	Suberitida	Suberitidae	*Aaptos*	*suberitoides*	*Aaptos suberitoides*	(Brøndsted, 1934)

Porifera	Homoscleromorpha	Homosclerophorida	Oscarellidae	*Oscarella*	*tuberculata*	*Oscarella tuberculata*	(Schmidt, 1868)
Sipuncula	Sipunculidea	Golfingiida	Golfingiidae	*Nephasoma*	*rimicola*	*Nephasoma (Nephasoma) rimicola*	(Gibbs, 1973)

Table S.5.2. Number of sequences generated in high-throughput sequencing and retaining along processing steps of bioinformatics pipeline for each primer-pair (COI - mlCOIintF/LoboR1; 18S - TAReuk454FWD1/TAReukREV3) and sampling times. RV – Ría de Vigo; RF – Ría de Ferrol. M – Mobile fauna sample; S – Sessile fauna sample. 6M – 6 months; 9M – 9months; 12M – 12 months. w/a – samples without amplification success.

Site	Substrate	Sampling time	Primer	Merged reads		Representative sequences		Not attributed*		Taxonomic assignment	
				M	S	M	S	M	S	M	S
RV	ARMS	6M	COI	42817	w/a	35675	w/a	2000	w/a	31979	w/a
			18S	38376	37178	32656	33128	321	149	1249	30465
		9M	COI	45397	w/a	37386	w/a	713	w/a	33121	w/a
			18S	34197	26970	30680	19234	245	186	15201	6083
		12M	COI	35476	51949	27476	36677	530	736	10889	23460
			18S	40793	38983	35907	34463	84	205	4065	10984
	ASMS	6M	COI	w/a	87606	w/a	62509	w/a	420	w/a	40521

			18S	25307	44730	22876	36230	13	222	14200	12142
		9M	COI	47424	72303	38104	49286	647	455	33250	35487
			18S	21058	31701	14497	28281	267	129	3934	5541
		12M	COI	42920	62934	34514	48608	651	148	13345	22120
			18S	36402	39885	27400	32643	258	120	5568	9454
RF	**ARMS**	**6M**	COI	w/a	w/a	w/a	w/a	w/a	w/a	w/a	w/a
			18S	38496	21980	22380	1539	1068	64	1734	937
		9M	COI	54568	46839	17682	37204	977	17	15456	12712
			18S	28859	37055	25599	4996	8	105	8930	1470
		12M	COI	38037	51638	31825	40437	130	122	25575	30951
			18S	47986	44838	43021	31732	8573	899	11968	26405
	ASMS	**6M**	COI	34835	42831	15189	30062	647	95	12929	8054
			18S	28690	44852	23969	23170	36	174	2446	11827
		9M	COI	30553	68684	23172	48087	111	467	15153	2821
			18S	44717	50118	35998	30679	182	1032	7056	16959
		12M	COI	33710	64169	23051	52180	154	77	15467	23121
			18S	40046	47508	35732	32395	10	125	27440	19109

*Not-attributed – singletons, rare sequences (<8 sequences) and non-metazoan assignments.

Table S.5.3. Average Taxonomic Distinctiveness (AvTD) based on the taxonomic information provided for the marine macrozoobenthic species detected in ARMS and ASMS in both sampling sites and among sampling times. 6M – 6 months; 9M – 9months; 12M – 12 months.

Sampling site	Substrate	Sampling time	AvTD (%)
RV	ARMS	6M	95.23
		9M	95.32
		12M	95.93
	ASMS	6M	94.57
		9M	94.08
		12M	94.38
RF	ARMS	6M	96.08
		9M	95.96
		12M	96.34
	ASMS	6M	95.62
		9M	96.02
		12M	96.33

Table S.5.4. SIMPER analysis showing species-level differences in community composition of the total species detected in each sampling site within sampling times and between sampling sites. 6M – 6 months; 9M – 9months; 12M – 12 months.

Sampling site	Sampling time	Average similarity within sampling time	Sampling times comparison	Average dissimilarity between sampling times
RV	6M	8.57	6M RV / 6M RF	79.59
	9M	32.00		
	12M	42.19	9M RV / 9M RF	71.21
RF	6M	21.62		
	9M	39.02	12M RV / 12M RF	70.31
	12M	36.17		

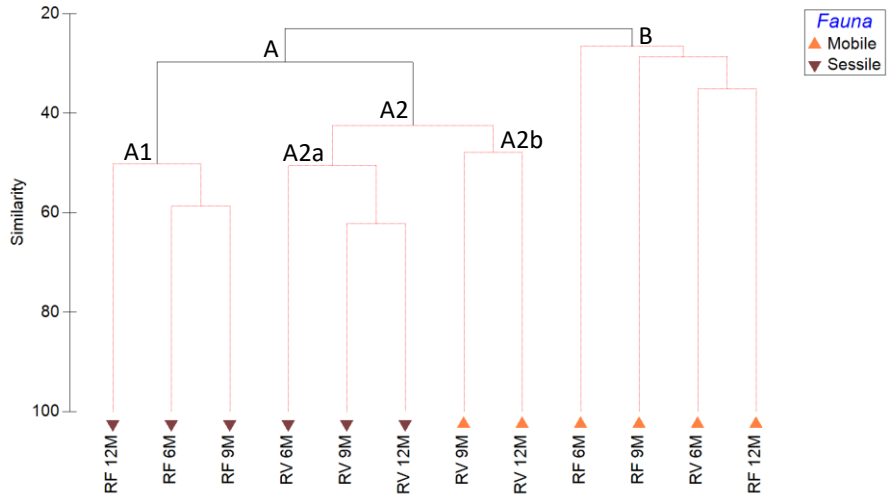


Figure S.5.1. Hierarchical clustering with SIMPROF tests (red bars), based on Bray-Curtis similarity index, and of the species detected in both substrates for mobile and sessile fauna fractions. Samples combine all the species detected in each sampling site (RV – Ría de Vigo and RF – Ría de Ferrol) among sampling times (6M – 6 months, 9M – 9 months and 12M – 12 months).

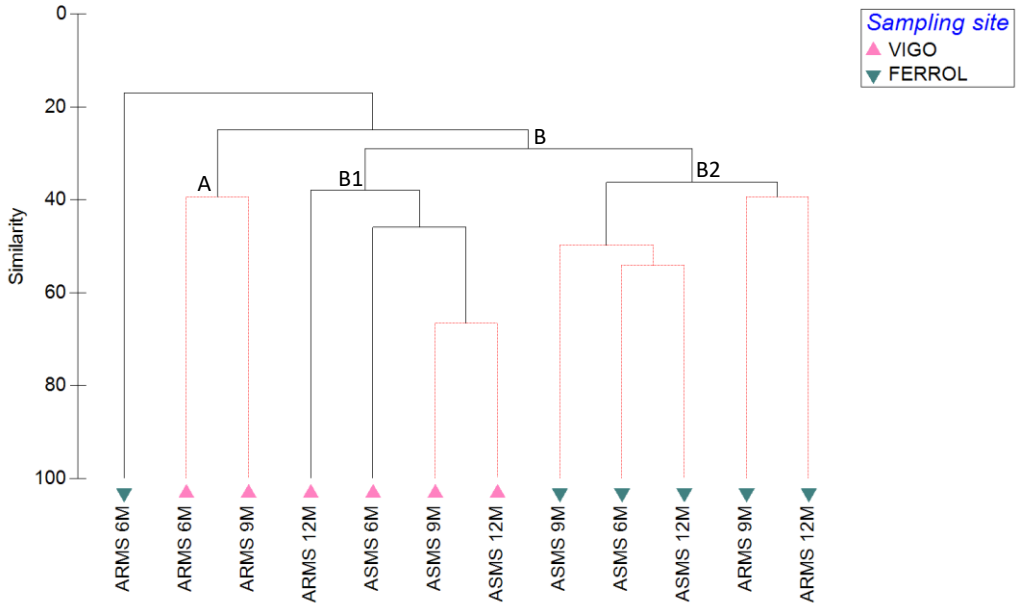


Figure S.5.2. Hierarchical clustering with SIMPROF tests (red bars), based on Bray-Curtis similarity index, of the species detected in ARMS and ASMS substrates. Samples combine the total species detected in each sampling site (VIGO – Ría de Vigo and FERROL – Ría de Ferrol) among sampling times (6M – 6 months, 9M – 9 months and 12M – 12 months).