



**HAL**  
open science

# Metabarcoding Is Powerful yet Still Blind: A Comparative Analysis of Morphological and Molecular Surveys of Seagrass Communities

Dominique A. Cowart, Miguel Pinheiro, Olivier Mouchel, Marion Maguer, Jacques Grall, Jacques Miné, Sophie Arnaud-Haond

► **To cite this version:**

Dominique A. Cowart, Miguel Pinheiro, Olivier Mouchel, Marion Maguer, Jacques Grall, et al.. Metabarcoding Is Powerful yet Still Blind: A Comparative Analysis of Morphological and Molecular Surveys of Seagrass Communities. PLoS ONE, 2015, 10 (2), pp.e0117562. 10.1371/journal.pone.0117562 . hal-01840066

**HAL Id: hal-01840066**

**<https://hal.umontpellier.fr/hal-01840066>**

Submitted on 26 May 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

RESEARCH ARTICLE

# Metabarcoding Is Powerful yet Still Blind: A Comparative Analysis of Morphological and Molecular Surveys of Seagrass Communities

Dominique A. Cowart<sup>1\*</sup>, Miguel Pinheiro<sup>2</sup>, Olivier Mouchel<sup>1</sup>, Marion Maguer<sup>3</sup>, Jacques Grall<sup>3</sup>, Jacques Miné<sup>4</sup>, Sophie Arnaud-Haond<sup>1\*</sup>

**1** IFREMER (Institut Français de Recherche pour l'Exploitation de la MER), Unité Environnement Profond, Département des Ressources physiques et Ecosystèmes de Fond de mer (REM), B.P. 70, 29280, Plouzané, France, **2** University of St. Andrews, Medical and Biological Sciences Building, North Haugh, St. Andrews, Fife, KY16 9TF, United Kingdom, **3** Institut Universitaire Européen de la Mer (IUEM), Technopôle Brest-Iroise rue Dumont d'Urville, 29280, Plouzané, France, **4** Total Exploration & Production, Direction HSE, 2 Place Jean Millier, 92078, Paris la Défense, France

\* [sophie.arnaud@ifremer.fr](mailto:sophie.arnaud@ifremer.fr) (SAH); [dominique.cowart@gmail.com](mailto:dominique.cowart@gmail.com) (DAC)



OPEN ACCESS

**Citation:** Cowart DA, Pinheiro M, Mouchel O, Maguer M, Grall J, Miné J, et al. (2015) Metabarcoding Is Powerful yet Still Blind: A Comparative Analysis of Morphological and Molecular Surveys of Seagrass Communities. PLoS ONE 10(2): e0117562. doi:10.1371/journal.pone.0117562

**Academic Editor:** Silvia Mazzuca, Università della Calabria, ITALY

**Received:** June 6, 2014

**Accepted:** December 27, 2014

**Published:** February 10, 2015

**Copyright:** © 2015 Cowart et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Sanger sequence data are available in the NCBI GenBank database (accession numbers KJ182970 – KJ183017).

**Funding:** This research was funded by a grant from Total Exploration & Production (FR00003996) to SAH. The funding institution had no role in the data collections, analyses and interpretations of this study. However, the funding organization did have a role in conceiving and designing the study, as well as in the contributions to the final manuscript. Additionally, the

## Abstract

In the context of the sixth wave of extinction, reliable surveys of biodiversity are increasingly needed to infer the cause and consequences of species and community declines, identify early warning indicators of tipping points, and provide reliable impact assessments before engaging in activities with potential environmental hazards. DNA metabarcoding has emerged as having potential to provide speedy assessment of community structure from environmental samples. Here we tested the reliability of metabarcoding by comparing morphological and molecular inventories of invertebrate communities associated with seagrasses through estimates of alpha and beta diversity, as well as the identification of the most abundant taxa. Sediment samples were collected from six *Zostera marina* seagrass meadows across Brittany, France. Metabarcoding surveys were performed using both mitochondrial (Cytochrome Oxidase I) and nuclear (small subunit 18S ribosomal RNA) markers, and compared to morphological inventories compiled by a long-term benthic monitoring network. A sampling strategy was defined to enhance performance and accuracy of results by preventing the dominance of larger animals, boosting statistical support through replicates, and using two genes to compensate for taxonomic biases. Molecular barcodes proved powerful by revealing a remarkable level of diversity that vastly exceeded the morphological survey, while both surveys identified congruent differentiation of the meadows. However, despite the addition of individual barcodes of common species into taxonomic reference databases, the retrieval of only 36% of these species suggest that the remaining were either not present in the molecular samples or not detected by the molecular screening. This finding exemplifies the necessity of comprehensive and well-curated taxonomic reference libraries and multi-gene surveys. Overall, results offer methodological guidelines and support for metabarcoding as a powerful and repeatable method of characterizing communities, while also presenting suggestions for improvement, including implementation of pilot

funding organization in no way restricted the data and results presented in this manuscript.

**Competing Interests:** The authors declare Total Exploration and Production as the funder of this study and the affiliation of JM, and this does not alter the authors' adherence to PLOS ONE policies on sharing data and materials.

studies prior to performing full “blind” metabarcoding assessments to optimize sampling and amplification protocols.

## Introduction

The sixth wave of extinction has already begun, far in advance of the completion of comprehensive biodiversity inventories [1,2]. Awareness of this situation has led to the establishment of transnational conservation programs whose effectiveness rely upon the ability to thoroughly assess biodiversity and provide indicators of ecosystem health within a time frame that counteracts the initial delayed response [3]. The construction of biological inventories has traditionally and primarily relied upon morphological identifications of taxonomic groups, however, morphological discrimination of a given community is a time consuming task that requires meticulous taxonomic expertise that is unfortunately becoming more rare [4,5]. Thus, there is a need for methods that can rapidly and cost effectively appraise ecosystem biodiversity and temporal variations following natural or impacted trajectories [4,6].

The improvement of molecular techniques in recent years have allowed for the development of genetic methods that help increase the rate and accuracy of species identification, at even the most remote ecosystems [7,8]. DNA barcoding, an approach in which target DNA sequences provide accelerated taxonomic identification, discrimination and discovery of unknown organisms, is at the forefront of these investigations [4,9,10]. DNA barcoding has become especially useful for analyses of environmental collections (water, soil, mud, feces, etc. . .), for which ‘environmental metagenetics’ is implemented when community sorting and morphological descriptions are challenging due to the large number and small sizes of possible taxon, as well as to the state of conserved specimens in samples [11–16].

Identifying early warning indicators of tipping points in ecosystems requires the most comprehensive appraisal of community biodiversity, as species or assemblages able to reveal the approach of critical thresholds may be invisible, cryptic or rare [17–20]. Ecosystem assessments performed by large-scale transnational conservation planning programs such as Natura 2000 (<http://www.natura.org/>), as well as by private companies, are also increasingly required to develop reliable environmental impact assessments (EIAs) before engaging in new activities. As a result of these requirements, “blind metabarcoding”, or ascribing taxonomic identity directly to sequences, has emerged as a possibly optimal solution for biodiversity surveys and inventories, in terms of costs and time schedules [21,22].

Morphological and molecular approaches have long been considered complementary [23,24], raising legitimate doubts as to the accuracy and reliability of blind metabarcoding. Thus far, several studies have shown the increased time efficiency of molecular techniques [25,26], particularly for specific taxa (fishes: [27], nematods: [28,29], arthropods: [24]). Additionally, several studies have addressed unknown and unrecognizable taxa [7], stomach contents [27,30,31] or microbial diversity [32]. The reliability of “blind metabarcoding” in those studies was tested through the analysis of controlled laboratory admixture or plausibility of taxa uncovered, whereas very few studies have rigorously compared the efficiency of molecular versus *in situ* morphological community descriptions from the same areas sampled for molecular analysis (*see* [24]). Single studies have, however, demonstrated great advances in estimating gains in terms of time saved and biodiversity revealed, by focusing on a single target gene and specific taxonomic groups such as Cytochrome Oxidase I (COI) for birds and arthropods [24], and the small subunit 18S ribosomal RNA region (18S) for nematodes and zooplankton [6,29].

In the present study, we aimed to test the efficiency and reliability of metabarcoding for environmental survey by providing one of the first comprehensive and rigorous comparisons of morphological and molecular characterizations of invertebrate communities. We characterized communities associated with *Zostera marina* seagrass meadows, using both mitochondrial and nuclear ribosomal genetic markers (Cytochrome Oxidase I and the small subunit 18S ribosomal RNA region) to increase the proportion of informative fragments in environmental DNA extracts [33], as well as to compare the diversity estimates between two barcoding genes.

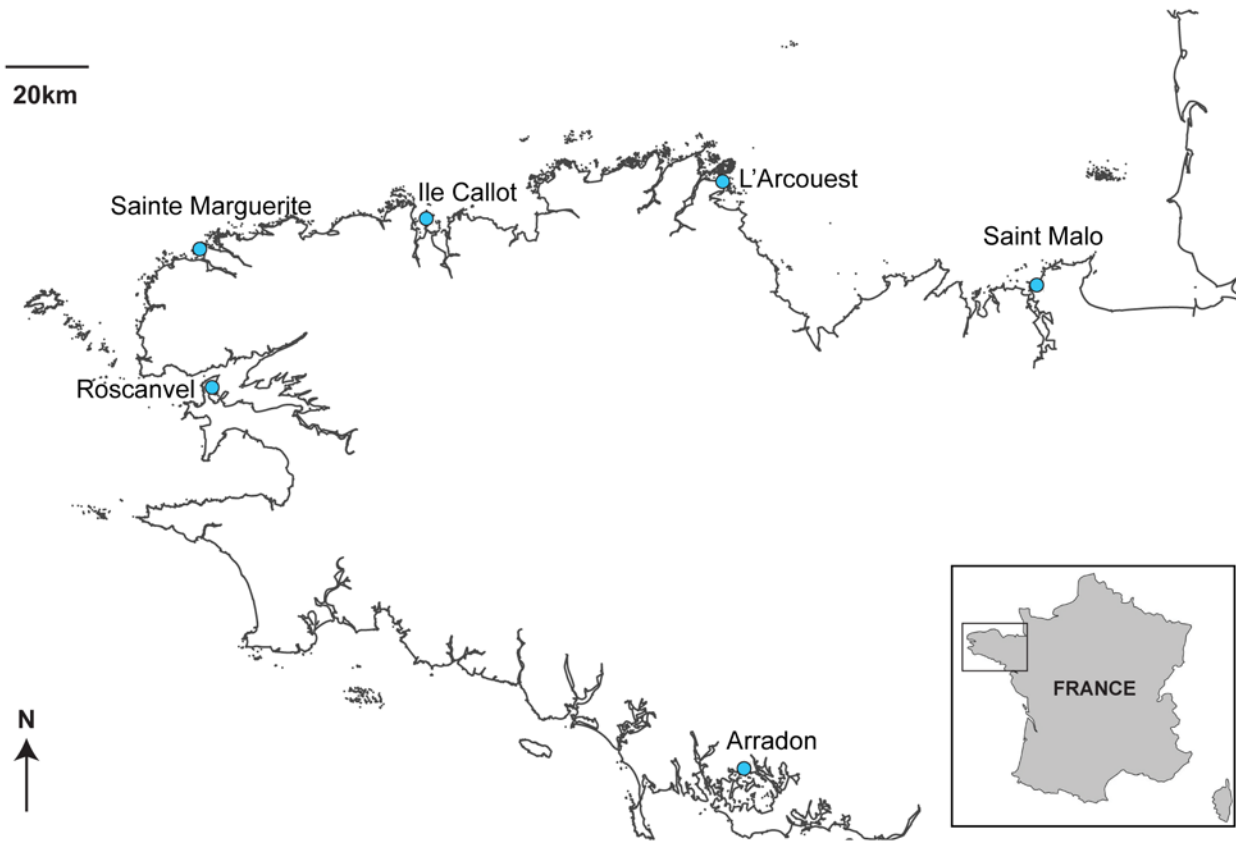
For alpha diversity estimates, the amount of taxonomic units revealed through a single molecular snapshot largely exceeded those uncovered through the morphological survey. Furthermore, the use of multi-gene metabarcoding also provided more reliable estimates of community composition than single genes, as each molecular marker has taxonomic specific affinities. Results also disclosed congruent patterns of beta diversity estimates across molecular and morphological surveys to confirm the reliability of metabarcoding, in addition to the molecular surveys uncovering meiofauna taxa. We also identified pitfalls and suggestions for improvement of molecular inventories, including the need for multi-marker assessments to unravel the broadest possible taxonomic diversity, as well as improved reference databases. Simultaneously, this comprehensive and comparative approach has allowed us to define an optimal sampling design that includes triplicate core samples for testing statistical significance, size trimming to limit the dominance of larger organisms, and a preliminary combination of broad taxonomic spectrum DNA markers to optimize estimates of biodiversity to be further used for ecological applications.

## Methods

### Sediment collections and sample processing

Sediment samples were collected from six *Zostera marina* seagrass meadows along the Brittany coast in western France (Fig. 1). No specific permissions were required to sample sediment at any of these locations, and the fieldwork performed here did not involve endangered or protected species. These meadows have been followed as part of an eight-year benthic survey, in which morphological identifications were performed by the observatory of the Institut Universitaire Européen de la Mer (IUEM), a component of the REseau BENThique (REBENT network [34]). The locations of each meadow are detailed in Table 1. The sampling protocol for the morphological characterization included two main collection methods: I) the collection of mobile epi-macrofauna, which were captured using pushing-nets on 10 m<sup>2</sup> surface, and II) the collection of soil-dwelling macrofauna, which were collected by sampling of sediments cores (0.03 m<sup>2</sup>). For the molecular characterization each location, two 20 X 30 m quadrats, that were previously sampled for the morphological survey and spaced several meters apart from one another, were chosen for sampling [35]. Three sediment cores per quadrat were collected during spring high tide at Sainte Marguerite and Arradon in 2010, and Sainte Marguerite, Ile Callot, L'Arcouest, Roscanvel and Saint-Malo in 2011, with the exception of Arradon, where only two cores could be analyzed (Table 1).

The exact same quadrates were rigorously chosen and sampled at similar dates for both surveys, yet morphological surveys require net and core sampling of extensive spatial areas, leading to an amount of material (tissues and sediment) not realistic to process for DNA extraction, which requires a smaller amount of homogenized material. To this end, sediment cores measuring 10cm diameter by 15cm depth were sieved on site using local seawater through decreasing mesh sizes of 2mm, 1mm, and 0.5mm in tandem order. This was done to identify if filtering by size uncovers additional metazoan diversity, as well as to prevent the dominance of larger bodied animals (see Fig. 2 for schematic of sampling protocol). As three mesh sizes were analyzed for each core sampled from each meadow, the term 'sample' refers to the sediment of a specific mesh size. More specifically, as the sediment remaining in a specific



**Fig 1. Map of six *Zostera marina* seagrass meadows along the coast of Brittany, France, where sediment collections were performed.**

doi:10.1371/journal.pone.0117562.g001

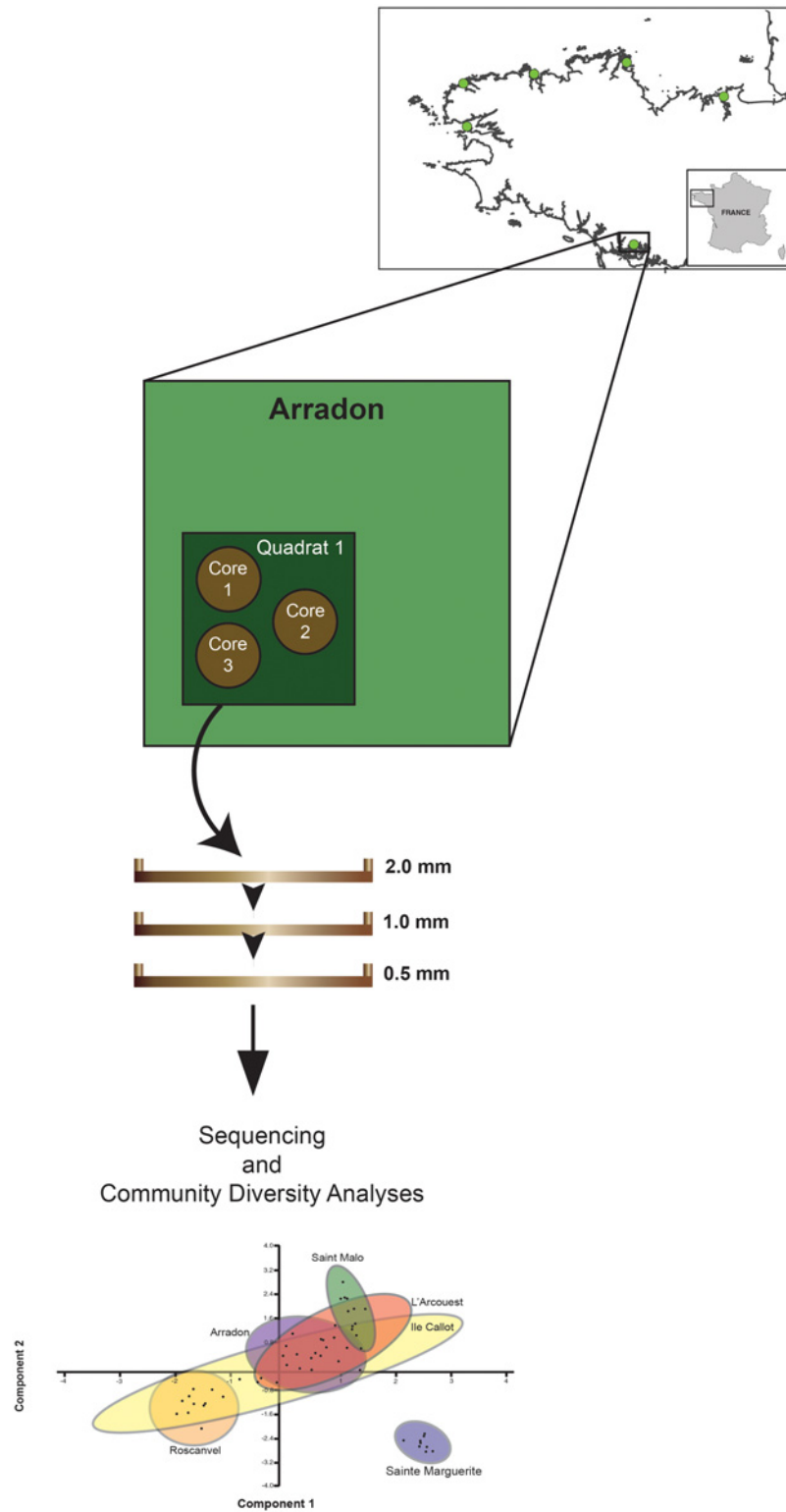
filter was collected as an individual sample, there were three samples collected for each core, two to three cores collected per quadrat, and two quadrats per meadow. Sieved sediment samples were immediately preserved at  $-80^{\circ}\text{C}$  in zip-locked bags until DNA extractions were performed on about 10g of each sample (i.e. each sieved fraction of each core), according to the protocol of Power<sup>TM</sup> Soil DNA Extraction Kit (MO BIO Laboratory, Solana Beach kits CA, USA). DNA extract quality was assessed using NanoDrop (Thermo Scientific).

Amplifications of two common barcoding genes, Cytochrome Oxidase (COI) and the small subunit 18S ribosomal RNA region (18S), were performed using “universal” primers for COI [36] and 18S [37], obtaining fragments of about 710bp and 450bp, respectively. Degenerate COI primers (F\_dgLCO-1490 5'-GGT CAA CAA ATC ATA AAG AYA TYG G and R\_dgHCO-

**Table 1. Locations of six *Zostera marina* seagrass meadows from where sediment collections were made.** Implemented barcoding markers were Cytochrome Oxidase (COI) and the small subunit 18S ribosomal RNA region (18S).

Meadow	Latitude	Longitude	Year(s)	No. of quadrats	No. of cores	Mesh sizes (mm)	Genes
Saint Malo	48.648° N	2.007° W	2011	2	3	0.5, 1.0, 2.0	COI
L'Arcouest	48.818° N	3.008° W	2011	2	3	0.5, 1.0, 2.0	COI
Ile Callot	48.697° N	3.925° W	2011	2	3	0.5, 1.0, 2.0	COI
Sainte Marguerite	48.596° N	4.623° W	2010/2011	1/2	3/3	0.5, 1.0, 2.0	18S & COI
Roscanvel	48.317° N	4.547° W	2011	2	3	0.5, 1.0, 2.0	COI
Arradon	47.626° N	2.822° W	2010	1	2	0.5, 1.0, 2.0	18S & COI

doi:10.1371/journal.pone.0117562.t001



**Fig 2. Sediment sampling schematic for the present study.**

doi:10.1371/journal.pone.0117562.g002

2198 5'-TAA ACT TCA GGG TGA CCA AAR AAY CA-3, [38]) were found to be suitable for the universal amplification of the COI fragment by revealing more diversity than the non-degenerated primers in a test of 454 sequencing, and were therefore used to amplify the COI fragment in 105 samples. For 15 of those samples from two sites, Sainte-Marguerite and Arradon (see Table 1), we additionally amplified 18S using the same primers as those previously used for metabarcoding of marine sediments and water samples [6,25,26]. As COI is widely known as the standard barcoding molecule for its ability to discriminate species in many animal groups [9,39] and in addition to Tang and others [5] identifying COI as uncovering more meiofaunal diversity than 18S, we chose to focus on COI for the large scale comparison of morphologically vs. molecularly characterized communities for all six seagrass meadows.

Two PCR replicates were performed for each sample using PolymeraseTwo—Platinum Taq High Fidelity (Invitrogen, CA, USA) from 50 ng of genomic DNA. Amplification was performed using the following PCR conditions: 95°C (2 min), 35 cycles of 95°C (1 min), 57°C (45 s), 68°C (3 min), followed by an elongation step of 68°C for 10 min. PCR products were visualized on 1.5% agarose gels stained with ethidium bromide to confirm the presence of COI and 18S fragments. PCR purification was completed using “Ampure beads” (Beckman Coulter, MA, USA). Amplification products were quantified by fluorimetry using the Invitrogen “PicoGreen” kit (CA, USA), and clustered in equimolar concentrations for sequencing reactions with 454 GS FLX Titanium, as instructed by the manufacturer (Roche, 454 Life Sciences, Branford, CT, USA). Unique sequences of 8bp were used as identifier tags to differentiate pooled samples and sequencing primers during bioinformatics analyses.

## Sequencing of REBENT taxonomically identified metazoans

To determine if commonly observed and morphologically identified species could be uncovered in our molecular datasets, tissues from the 50 most common benthic metazoan species from *Zostera marina* seagrass communities (Table 2) were isolated for DNA extraction and sequencing. The 50 species were collected and identified as a part of the morphological inventory compiled by the IUEM/REBENT network [34], and classified as the most common due to their consistent presence and high abundance at the six seagrass meadows described in this study. A detailed protocol for the extraction and sequencing of these metazoan species can be found in the S1 Protocol in the supplemental material section. A 710bp fragment of COI was successfully amplified and sequenced in 14 of the 50 species, while a 450bp fragment of 18S was successfully amplified and sequenced for 33 of those species, each using the respective universal primer set. As public databases may contain sequence data that has not been taxonomically curated, the isolated sequences we amplified (known as common species “barcodes”) were added to the reference databases used for taxonomic assignments, to increase the accuracy of the assignments.

## Bioinformatics data processing and analyses

Raw pyrosequenced reads for COI and 18S were processed, clustered and taxonomically assigned, using the Quantitative Insights into Microbial Ecology (QIIME) v. 1.7.0 pipeline [40]. A workflow of QIIME scripts executed for this project can be found in S2 Protocol of the supplemental material section. During the processing stage, dataset demultiplexing and quality checks, including the removal of low quality and short sequences (< 200bp), were implemented using the `split_libraries.py` script. Filtered reads were clustered into *de novo* molecular operational taxonomic units (MOTUs, [41]) with the aid of UCLUST [42], using the script `pick_otus.py` in QIIME at a defined pairwise sequence identity cut off value was 97% [5,6,43]. Next, a representative sequence set of the clusters was generated using the `pick_rep_set.py`

**Table 2. The 50 most common invertebrate metazoan found in *Zostera* seagrass meadow sediment (REBENT, 2010).** x denotes taxon found in NCBI GenBank, \* denotes taxa barcode added to NCBI GenBank (this study), F (Family), G (Genus) and S (Species) denote taxonomic level found in the COI and 18S environmental sequenced datasets for this study.

Phylum	Class	Family	Species	COI Genera	COI Species	Presence in COI Metadataset	18S Genera	18S Species	Presence in 18S Metadataset
<b>Annelida</b>	Polychaeta	Ampharetidae	Melinna palmata	*	*		x	*	S
		Capitellidae	Notomastus latericeus	x			x	x	S
		Capitellidae	Heteromastus filiformis				x	x	F
		Capitellidae	Capitella capitata	x	x		x	x	G
		Cirratulidae	Caulleriella bioculata				x		F
		Cirratulidae	Cirriformia tentaculata				x	x	F
		Cirratulidae	Chaetozone setosa	x	x	S			F
		Glyceridae	Glycera alba	x			x	x	
		Lumbrineridae	Scoletoma impatiens	x					G
		Lumbrineridae	Lumbrineris latreilli	x	x		x	x	G
		Maldanidae	Euclymene oerstedii	x	*	S	x	x	S
		Maldanidae	Clymenura clypeata	*	*	S	x	x	S
		Nephtyidae	Nephtys hombergii	x	x	S	x	x	S
		Nereididae	Platynereis dumerilii	x			x	x	S
		Nereididae	Neanthes caudata	x			x	*	S
		Nereididae	Perinereis cultrifera	x			x	*	S
		Orbiniidae	Scoloplos armiger	x	x		x	x	S
		Phyllodoceidae	Phyllodoce mucosa	x	x	F	x		F
		Phyllodoceidae	Eteone longa	x	x	F	x	x	G
		Sabellidae	Megalomma vesiculosum	x			x	*	S
Sigalionidae	Sthenelais boa	*	*	S	x	x			
Spionidae	Aonides oxycephala				x	x	S		
Spionidae	Spio martinensis	x			x		G		
<b>Arthropoda</b>	Malacostracea	Ampeliscidae	Ampelisca brevicornis	x	*		x	*	
		Ampithoidae	Ampithoe rubricata	x	x	S	x	x	
		Aoridae	Microdeutopus anomalus					*	
		Apseudidae	Apseudes latreilli	x			x	*	S
Apseudidae	Apseudes talpa	x	x		x	*	G		

(Continued)



Table 2. (Continued)

Phylum	Class	Family	Species	COI Genera	COI Species	Presence in COI Metadataset	18S Genera	18S Species	Presence in 18S Metadataset
		Atylidae	<i>Atylus guttatus</i>	*	*		x	x	
		Atylidae	<i>Atylus swammerdami</i>	*	*		x	*	
		Bodotriidae	<i>Iphinoe trispinosa</i>	x	x		*	*	
		Caprellidae	<i>Phtisica marina</i>				x	x	
		Caprellidae	<i>Caprella acanthifera</i>	x			x	*	
		Dexaminidae	<i>Dexamine spinosa</i>	*	*		*	*	
		Hippolytidae	<i>Hippolyte varians</i>				x	x	
		Lysianassidae	<i>Lysianassa plumosa</i>	x		F			
		Melitidae	<i>Gammarella fucicola</i>				*	*	
		Paguridae	<i>Anapagurus hyndmanni</i>	*	*		*	*	
		Portunidae	<i>Polybius arcuatus</i>	x					
		Tanaidae	<i>Tanais dulongii</i>	x	x		x	*	
		Urothoidae	<i>Urothoe pulchella</i>	*	*		x	*	
Cnidaria	Anthozoa	Actiniidae	<i>Anemonia viridis</i>		x		x	x	G
Mollusca	Bivalvia	Lucinidae	<i>Lucinoma borealis</i>				x	x	S
		Montacutidae	<i>Mysella bidentata</i>	*	*		x	*	
		Semelidae	<i>Abra alba</i>	x			x	x	
	Gastropoda	Calyptreaeidae	<i>Calyptrea chinensis</i>	x	x	F			
		Nassariidae	<i>Nassarius reticulatus</i>	x	x		x	*	
		Trochidae	<i>Gibbula cineraria</i>	x	x	F	x	x	
		Trochidae	<i>Jujubinus striatus</i>	x	x	S	*	*	
Sipuncula	Sipunculidae	Golfingiidae	<i>Golfingia elongata</i>				x	x	S
<b>Total in NCBI public database</b>				<b>39</b>	<b>26</b>	<b>F(5), G(0), S(5)</b>	<b>44</b>	<b>41</b>	<b>F(5), G(7), S(14)</b>

doi:10.1371/journal.pone.0117562.t002

script. Chimeras and “quasi-singletons” (sequences appear < 3 times in the dataset) were identified and removed from the MOTU table using USEARCH v.6.1 [42] implemented in QIIME.

The 33 and 14 prescreened and taxonomically confirmed DNA barcode sequences for 18S and COI (accessions KJ182970—KJ183017) were added to previously compiled reference databases for taxonomic assignment. Fourteen (18S) and three (COI) of these prescreened and taxonomically confirmed sequences were already present in the GenBank public database ([www.ncbi.nlm.nih.gov/genbank](http://www.ncbi.nlm.nih.gov/genbank)), based only upon their shared sequence names (see Table 2). For 18S, the morphologically confirmed barcodes were added to the SILVA SSU r115 EMBL *eukarya* database ([44], <http://www.arb-silva.de/>) to produce a reference database composed of 11,617 sequences. The COI barcodes were added to a COI *metazoan*-only reference database that we compiled from GenBank to produce a database composed of 48,734 sequences. These databases were used to assign taxonomic identifications to representative MOTU sequences based on the highest similarity score using the BLAST assignment method [45] implemented

through the `assign_taxonomy.py` script in QIIME. Sequences that did not return significant hits at >90% identity were labeled as “no blast hit” (unassigned), and we attempted to determine their putative identity by using inferences based upon global blast search.

Molecular data were analyzed on the basis of presence/absence of taxonomically assigned MOTUs in each sample, as quantitative data based on the number of sequences ascribed to a given MOTU are unlikely to provide accurate quantitative estimates due to factors other than species abundance that likely interfere with the number of times a sequence was observed (e.g. number of biological cells present in the sediment, DNA density in tissue, primer matching, etc.). For the most accurate comparisons of metazoan diversity across morphological and molecular datasets, analyses were performed with and without the unassigned MOTUs. While the results were congruent, those obtained excluding the unassigned MOTUs are reported here, and results including these MOTUs can be found in the supplemental information section. The decision to exclude unassigned MOTUs was particularly important for the COI dataset, which had the most unassigned (93% of all MOTUs). We determined the putative identity of the unassigned COI and 18S MOTUs with the aid of NCBI’s BLAST stand-alone program [46], using the public nucleotide database implemented with the `blastn` algorithm at default parameters.

Presence/absence information obtained from the molecular datasets were compared to a quantitative morphological dataset containing 322 species collected and taxonomically identified by the IUEM/REBENT network in 2011, in order to appraise the loss or gain of information obtained with each method in terms of alpha and beta diversity estimates. Given the inherent difficulties of morphologically identifying organisms smaller than 1mm, the morphological dataset was generated from animals with body sizes  $\geq 1$ mm, while the molecular datasets included meiofauna (animals with body sizes  $< 1$ mm [47]) in addition to macrofauna. Therefore, diversity analyses were performed comparing 1) morphological data ( $\geq 1$ mm) with all molecular data ( $\geq 0.5$ mm) and 2) morphological data ( $\geq 1$ mm) with molecular data  $\geq 1$ mm (excluding 0.5mm) for consistency and to better evaluate similarities and differences across survey techniques. A list of the 322 morphologically identified species and their presence/absence in each meadow is shown in [S1 Dataset](#).

Multiple rarefactions were applied to the molecular datasets to obtain standardized estimates of alpha diversity using Chao1 measurement for species richness [48, 49]. Chao1 values were calculated using the `specpool` function of the Vegan community ecology package (v.2.0–10), executed in the statistical software R (v.3.0.2) [50]. Analysis of Similarities (ANOSIM) testing was done to determine if significant differences exist between sampling groups (core, mesh size, meadow), while Principal Component analyses (PCA) were completed to provide spatial illustrations of community structure across meadows, and was performed on the basis of Jaccard distances to minimize the weight given to absence (0) values. Both ANOSIM and PCA analyses were carried out using the software PAST [51]. Finally, taxonomic community compositions by phyla were illustrated via a phylogenetic tree produced using the Interactive Tree of Life iTOL tool [52].

## Results

### Molecular and morphological datasets

The Roche 454 FLX sequencing of *Zostera marina* sediment produced a total of 622,468 and 190,509 raw sequences, which were quality filtered to 412,838 and 153,463 for COI and 18S, respectively ([S1 Table](#)). The additional removal of singletons and chimeras produced 411,086 (COI) and 150,677 (18S) sequences, which were then clustered into 13,492 and 1,316 representative MOTUs for COI and 18S, respectively.

[Table 2](#) presents a list of the 50 most common metazoan species recovered and morphologically identified from the sediment of *Z. marina* seagrass by the IUEM for the REBENT

network. The list details the presence of COI and 18S barcode sequences matching the genera and/or species available in the public sequence database GenBank (NCBI), as well as those families, genera or species that match to MOTU sequences in the environmental metadatasets. At the present, 51% (COI) and 80% (18S) of sequence barcodes of the 50 most common and abundant metazoan species are publically available in NCBI, encompassing 35 families and 48 genera (Table 2). From the COI environmental metadataset, seven MOTUs blasted up to the species level (14% of all common species and 50% of the barcodes added to the COI reference database), and five to the family level. For the 18S metadataset, 14 MOTUs blasted up to the species level (28% of all common species and 42% of the barcodes added to the 18S reference database), seven to the genera and five to the family (Table 2).

## Uncovering the identity of taxonomically unassigned MOTUs

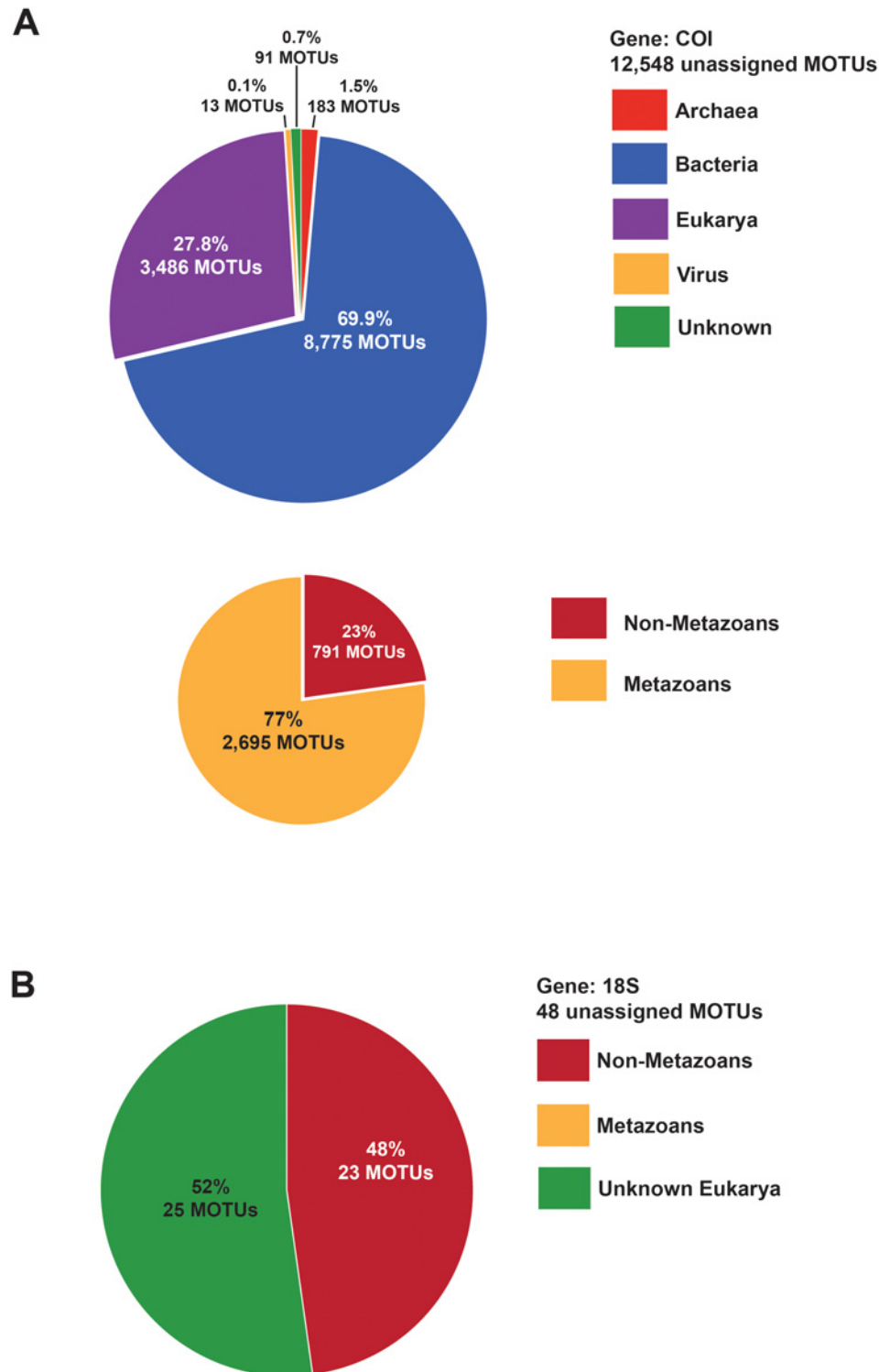
“No blast hit”, or unassigned matches, recognized an inability of the BLAST method implemented through QIIME to uncover a close sequence matching the target with an identity of >90%.

Of the 13,492 MOTUs uncovered for COI, 944 were assigned to metazoans, while 12,548 (93%) remained unassigned. To glean the identity of the unassigned MOTUs, we performed the *blastn* procedure for each unassigned sequence against NCBI’s global public database with the aid of the stand-alone BLAST command line program. The overall findings identified the closest match for 1.5% of the MOTU sequences as Archaea, 69.9% as Bacteria, 27.8% as Eukarya, 0.1% as viruses and 0.7% as still unknown/unidentified with no match in the public database (Fig. 3A, S2 Dataset). Further analysis of these 0.7% unknown revealed open reading frame interruptions, suggesting the amplification of pseudogenes, possibly “*numts*”, or nuclear copies of mitochondrial derived genes that become non-functional/coding and are often seen in COI [53, 54] (Fig. 3A). As the *blastn* algorithm can be less stringent than the BLAST program implemented in QIIME, we cannot discard the occurrence of *numts* among the MOTUs unassigned after QIIME BLAST but assigned after the *blastn* procedure. Therefore, while *numts* are expected to fall close to their original mitochondrial sequence, we detail only analyses with the 944 initially assigned MOTUs below. We did, however, compare diversity results including and excluding the initially unassigned 12,548 MOTUs (S2 Table, S3 Table, S4 Table, S5 Table, S1 Fig.). The results for both datasets showed similar qualitative results for community structure, although diversity estimates are lower bounded when excluding the initially unassigned MOTUs.

Of the 1,316 MOTUs uncovered for 18S, 1,174 were identified as metazoans and 94 as non-metazoans, while 48 (3.5%) remained unassigned, which is a relatively low percentage compared to what has been observed previously (10%, [55]). To glean the putative identity of the 48 unassigned MOTUs, these sequences were also blasted against the NCBI public database, for which we found 48% of the initially unassigned MOTUs matched to non-metazoan eukaryotes, while 52% were still unknown (Fig. 3B, S3 Dataset). Given that the 52% of unknown MOTUs could contain metazoans, we compared diversity results excluding and including unassigned 18S MOTUs, but for the rest of manuscript, we discuss only the MOTUs initially assigned, similarly to what was done for the COI dataset. Both diversity analyses including and excluding the 48 initially unassigned MOTUs also show similar results (S2 Table, S3 Table, S6 Table, S7 Table, S2 Fig., S3 Fig.).

## Taxonomic community composition by phyla

The number and frequency of MOTUs assigned to 13 major metazoan phyla uncovered by morphologic and molecular methods are illustrated in Fig. 4. Colored bar lengths correspond to the frequency of MOTUs assigned to a specific phyla by each survey method, while the numbers above each frequency bar refer to the number of MOTUs identified. Numbers in brackets [N]



**Fig 3. Taxonomic percentages and counts of Molecular Operational Taxonomic units (MOTUs) initially unassigned for COI and 18S, inferred using NCBI BLAST public nucleotide database.** (A) Taxonomic percentages for the 12,548 unassigned MOTUs for the COI gene. The top pie chart illustrates the proportions of each domain, while the bottom pie chart estimates the proportions of metazoan versus non-metazoans within the eukaryotes. (B) Taxonomic percentages for the 48 unassigned MOTUs for the 18S gene. The pie chart is divided into non-metazoans versus unknown Eukaryotes. The initially unassigned MOTUs for both genes were identified using the blast-stand alone program with the “blastn” algorithm under default parameters.

doi:10.1371/journal.pone.0117562.g003

above the COI bars correspond to the number of COI MOTUs that were initially unassigned, but once individually blasted to identify taxonomy, matched to one of the major phyla. Numbers without brackets above COI correspond to those MOTUs that were initially identified ( $n = 944$ ). All frequency values can be found in the [S4 Dataset](#). While Annelida, Arthropoda and Mollusca were the most frequently observed phyla with the highest number of MOTUs being assigned to these groups, each method differed in its taxa specific affinity. The COI survey identified arthropods as the most common phylum with and without taking into account the previously unassigned MOTUs, followed by mollusks, whereas 18S had the highest number of MOTUs assigned to annelids followed by arthropods, and revealed fewer molluscan taxa. Both morphological and COI surveys were better for unraveling Chordata and Echinodermata, whereas 18S identified Nematoda and Porifera (sponges). Taking into account previously unassigned COI MOTUs, COI was best at identifying Platyhelminthes (flatworms) and Cnidaria.

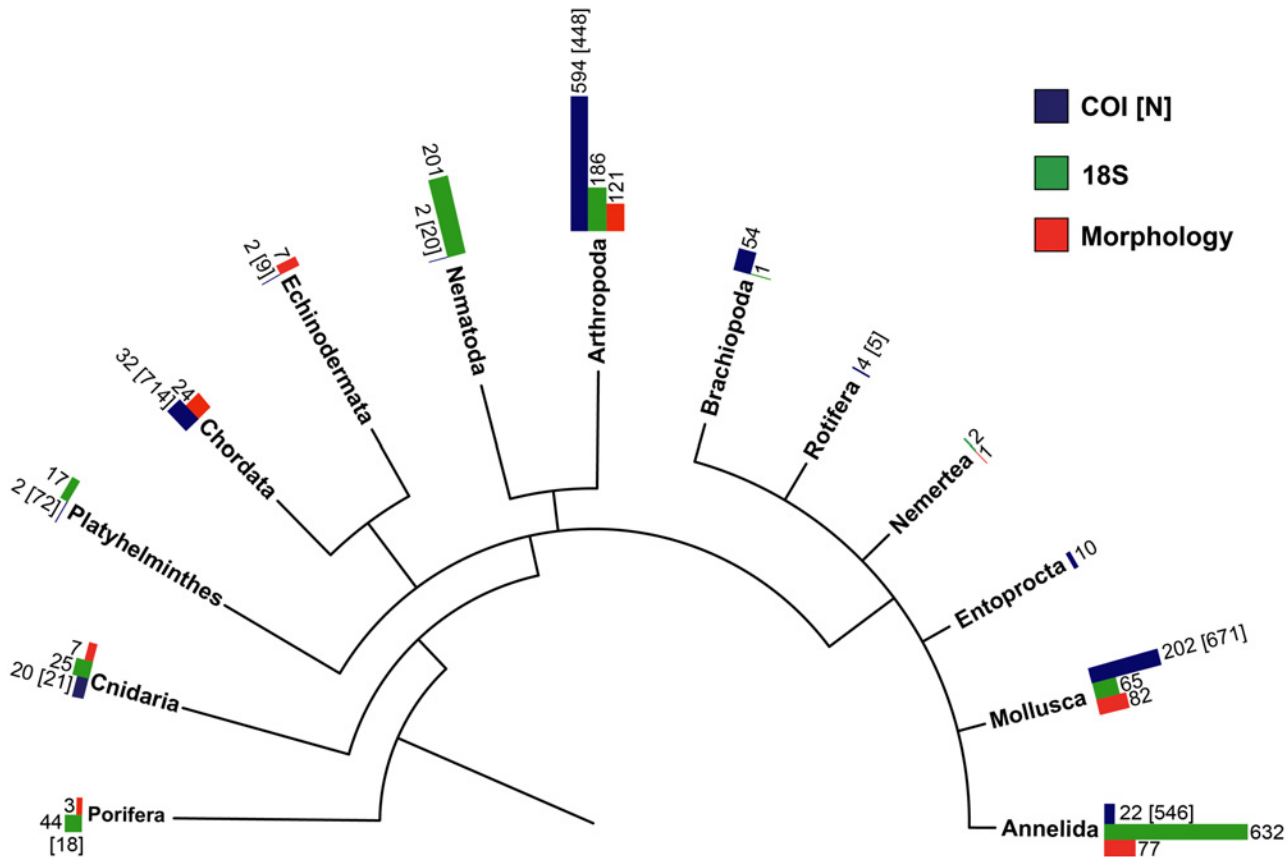
The frequencies of metazoan phyla uncovered in each of the three mesh size fractions are shown in [Fig. 5](#). The better capacity of COI for uncovering molluscs and arthropods and 18S for annelids and nematods is maintained, regardless of mesh size. For COI, annelids had the highest representation in 2mm at Arcouest and Arradon, while mollusks were detected at all size classes across meadows ([Fig. 5A](#)). Nematodes were barely detected using the COI survey, but an unusually high frequency of mollusks and chordates were found in the 0.5mm fraction at Saint Malo. The COI survey also identified mostly arthropods in the 1–2mm fractions of sediment at Sainte Marguerite in 2010 whereas a predominance of mollusks was revealed one year later in the same quadrat, a pattern that was consistent across core replicates.

For 18S, nematodes were dominant in the 0.5mm fraction, and their presence apparently decreases with larger size classes at both meadows analyzed ([Fig. 5B](#)). Arthropods were also highly represented in 0.5mm, while mollusks were mostly detected in the 2mm size class at Sainte Marguerite. At Arradon, the larger size classes appear heavily dominated by annelids. Lastly, the morphological survey consisted of  $\geq 1$ mm size fractions and identified a nearly even frequency of arthropods and mollusks at all meadows except the southernmost Arradon ([Fig. 5C](#)). At Arradon, representatives from Chordata and Nemertea appeared as frequent macrofauna, whereas mollusks were not detected. Overall, these results support the use of size filtering for unveiling an all-encompassing picture of community biodiversity.

## Alpha diversities and community comparisons across spatial scales

As the morphological survey consisted of only macrofauna, for the purposes of accurate comparisons, alpha diversity analyses were performed against molecular surveys including and excluding the 0.5mm fraction. Both molecular datasets consistently exhibited much higher values (mostly five to ten times higher) than that of morphology, and when including only assigned MOTUs for the purpose of being conservative, 18S exhibited much higher Chao1 values for Sainte Marguerite and Arradon compared to COI ([Table 3](#) and [Fig. 6](#), see [S5 Dataset](#), [S6 Dataset](#) and [S7 Dataset](#) for all diversity values). Additionally, comparisons of alpha diversity by meadow did not show strict correlation across morphological and molecular surveys ([Fig. 6](#), [S8 Table](#)), with meadows having the highest species richness values differing depending on the survey method implemented. In general, these analyses illustrated that even when excluding the 0.5mm fraction, molecular surveys deliver much higher diversities when compared to the morphological survey.

ANOSIM testing for differences in community structure across spatial scales revealed that for both morphological and molecular methods, meadows are significantly different from one another ( $p \leq 0.013$ ), while cores are not significantly different within meadows ( $p \geq 0.891$ , [Table 4](#)). For molecular analyses, mesh sizes also differed in terms of community composition (COI,  $p = 0.001$ , 18S,  $p = 0.014$ ), with the 0.5mm fraction being highly dissimilar from both of



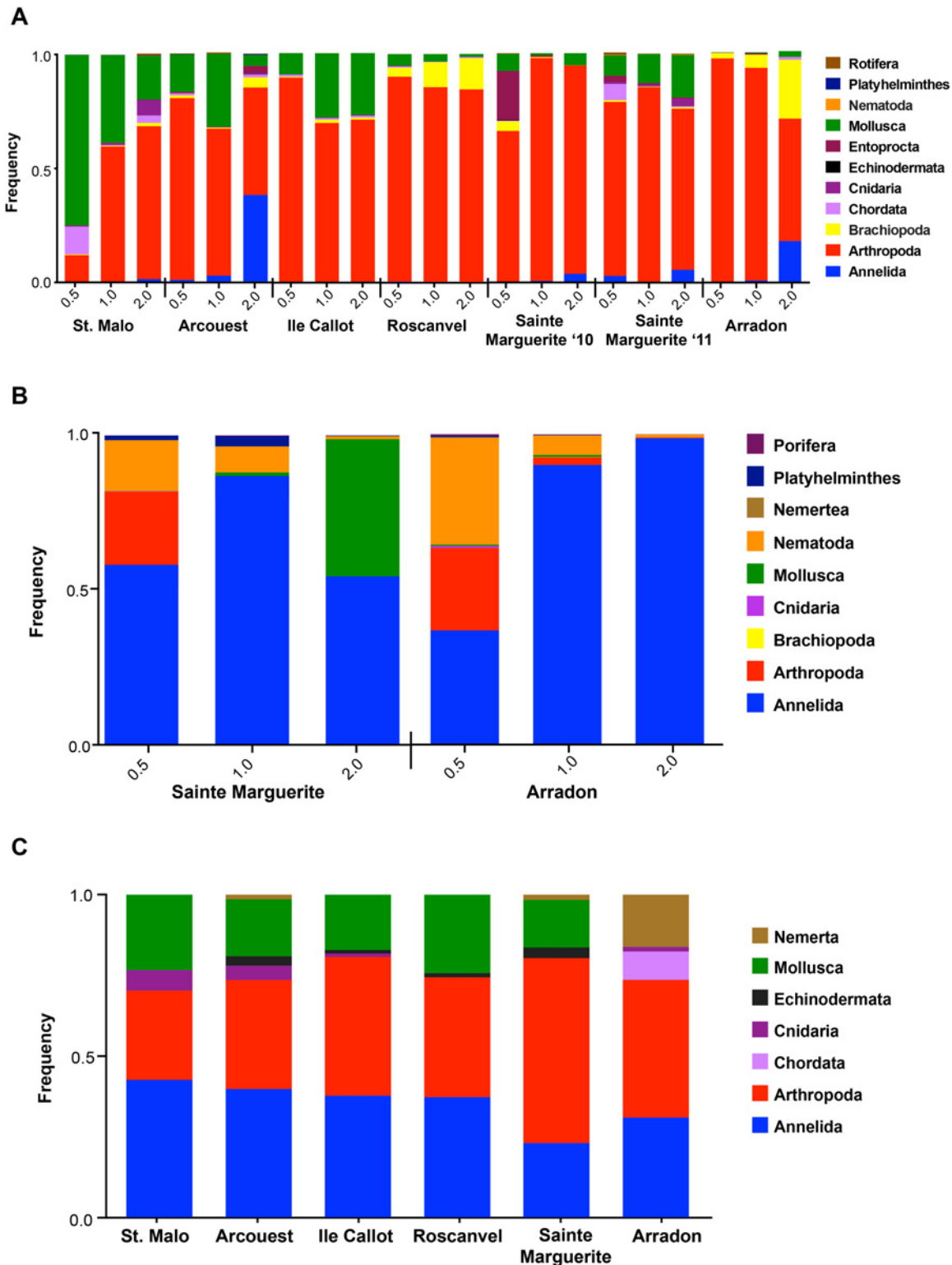
**Fig 4. Frequencies of phyla uncovered in *Zostera marina* sediment, using morphological and molecular (COI and 18S) surveys.** Bar lengths correspond to the frequency of molecular operational taxonomic units (MOTUs) assigned to a specific phyla for each survey method. The numbers above each frequency bar refer to the number of MOTUs identified by each respective survey method. Numbers in brackets [N] above the COI bars correspond to the number of COI MOTUs that were initially unassigned, but once individually blasted to identify taxonomy, matched to major phyla. Numbers without brackets above COI correspond to those MOTUs that were initially identified (n = 944). Tree was produced using the iTOL tool.

doi:10.1371/journal.pone.0117562.g004

the larger size classes (pairwise  $p \leq 0.03$ , [S4 Table](#), [S5 Table](#), [S6 Table](#), [S7 Table](#)). When results from the three mesh sizes were pooled to assess biodiversity present by core, cores were found to be homogeneous in terms of community composition within each meadow (COI,  $p = 0.945$ , 18S,  $p = 0.891$ , [S3 Table](#)). This result is in line with homogeneity also observed among morphologically described cores ([S9 Table](#) and [S10 Table](#)), supporting a lack of significant spatial heterogeneity that advocates the role of cores as replicates in the sampling scheme.

### Morphological and molecular multidimensional analyses

Principal Components Analyses (PCA) of each survey identified similar patterns of community composition ([Fig. 7](#)). The six meadows emerged as six distinct clusters supported by significant results from the ANOSIM analyses of both morphological and molecular (COI) surveys ([Table 4](#)), and the PCA analyses show the meadows arranged in a similar orientation for both morphological and molecular surveys ([Fig. 7](#)). Furthermore, Sainte Marguerite was defined by both methods as having a most dissimilar community composition when compared to the other meadows, which is likely the result of both oceanographic forces and demographic history of this specific meadow. Despite these similarities across surveys, mantel testing using



**Fig 5. Frequency comparisons of phyla found in three size class fractions 0.5, 1.0, and 2.0 mm.** (A) Frequency of phyla by size identified using the COI gene at six meadows. Analyses include two years (2010 and 2010) for the Sainte Marguerite meadow. (B) Frequency of phyla by size identified using the 18S gene at two meadows. (C) Frequency of phyla found at the  $\geq 1$  mm size class via morphological identifications.

doi:10.1371/journal.pone.0117562.g005

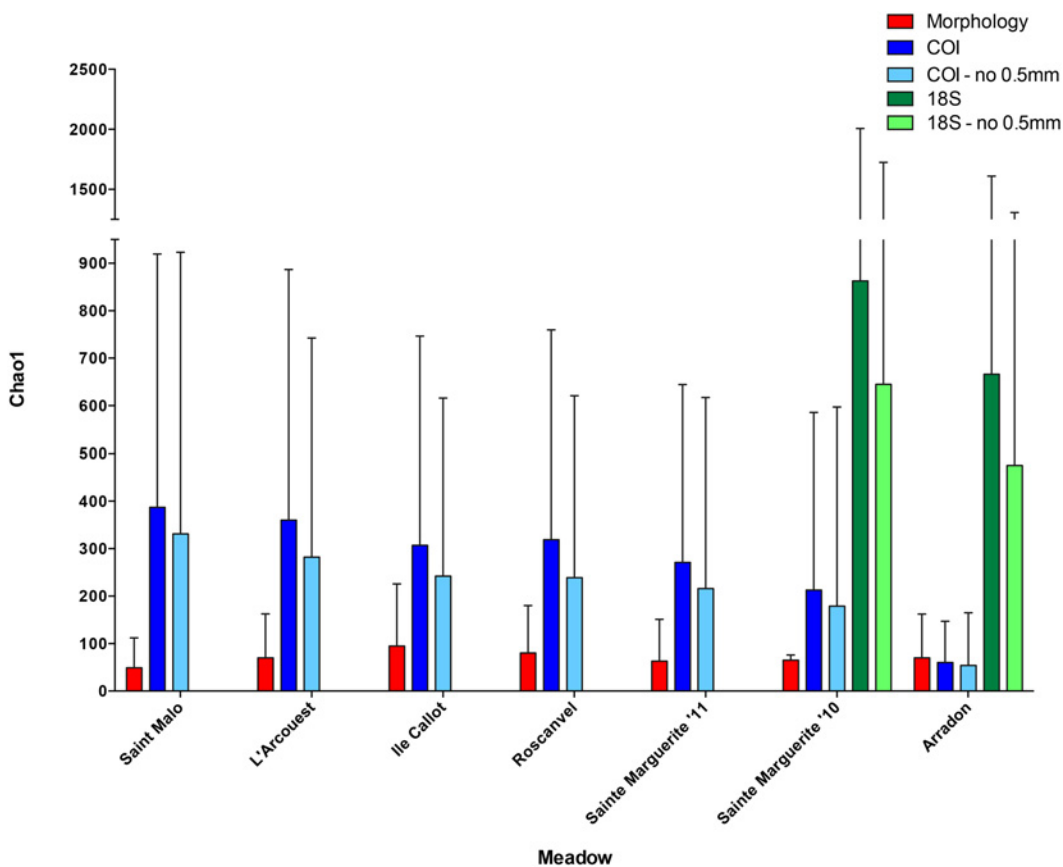
**Table 3. Chao1 values by *Z. marina* meadow for each survey method.** The value after the +/- symbol indicates the standard error.

Meadow	Method				
	Morphology	COI		18S	
			Including 0.5mm	Excluding 0.5mm	Including 0.5mm
Saint Malo	49 ± 63	387 ± 497.30	331 ± 555.12	—	—
L'Arcoouest	70 ± 92.78	360 ± 506.60	282 ± 474.53	—	—
Ile Callot	95 ± 131	307 ± 460.12	243 ± 351.93	—	—
Roscanvel	80 ± 100.34	319 ± 439.19	240 ± 339.01	—	—
Sainte Marguerite 2011	63 ± 63	271 ± 366.60	216 ± 351.84	—	—
Sainte Marguerite 2010	65 ± 11.22	213 ± 392.56	179 ± 417.34	863 ± 1144.43	645 ± 1075.27
Arradon	70 ± 92.22	60 ± 77.79	54 ± 116.23	666 ± 914.06	475 ± 806.2

doi:10.1371/journal.pone.0117562.t003

Jaccard beta diversity distances for comparison of morphological and molecular surveys showed non-significant results ( $p > 0.05$ , S8 Table).

Even with significant differentiation of the six clusters corresponding to each meadow via ANOSIM testing, the molecular COI PCA illustrated overlap between neighboring meadows of Roscanvel and Ile Callot, as well as Saint Malo and L'Arcoouest (Fig. 7B), similar to that



**Fig 6. Chao1 values presented by meadow for morphological and molecular surveys.** Dark bar colors for the molecular surveys denote the inclusion of three size fractions, while light bar colors denote the exclusion of the 0.5mm size fraction, done for accurate comparisons with morphological surveys. Chao1 values were calculated using Vegan community ecology package implemented in R [50].

doi:10.1371/journal.pone.0117562.g006



**Table 4. Analysis of similarities (ANOSIM) testing across meadow, core and mesh size spatial scales, performed in PAST [51].** NS = not significant, \* 0.5mm size fraction significantly differs from both 1 and 2mm fractions ( $p < 0.05$ ).

Taxonomic community assessment	Method		
	COI	18S	Morphology
Ho: meadows do not differ	$p = 0.001$	$p = 0.013$	$p = 0.001$
Ho: cores do not differ	NS	NS	NS
Ho: mesh sizes do not differ*	$p = 0.001$	$p = 0.014$	—

doi:10.1371/journal.pone.0117562.t004

obtained on the basis of the morphological survey. Stronger divergence of meadows was depicted with the COI survey, identifying samples clustering to their specific meadow, irrespective of core collection or mesh size.

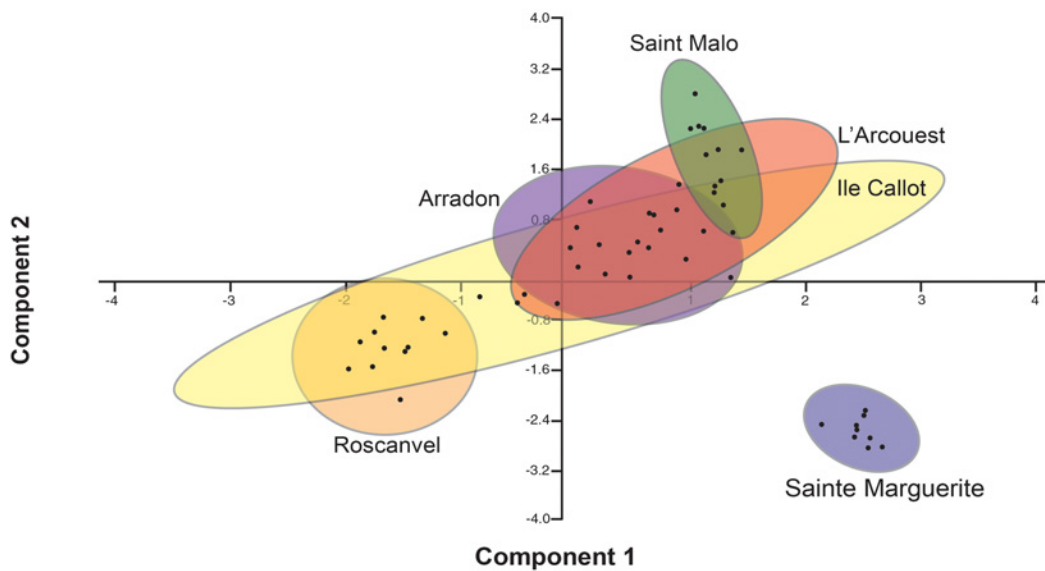
For the molecular survey of Sainte Marguerite, samples from 2010 (denoted by open circles within the cluster) and samples from 2011 (denoted by closed circles within the cluster) grouped together, suggesting that overall temporal community structure remained globally similar across 2010 and 2011 for this specific meadow. Finally, apart from Sainte Marguerite and Arradon, COI meadow community structure depicted an east/west gradient across the region beginning with Saint Malo, L'Arcouest, Roscanvel and Ile Callot.

## Discussion

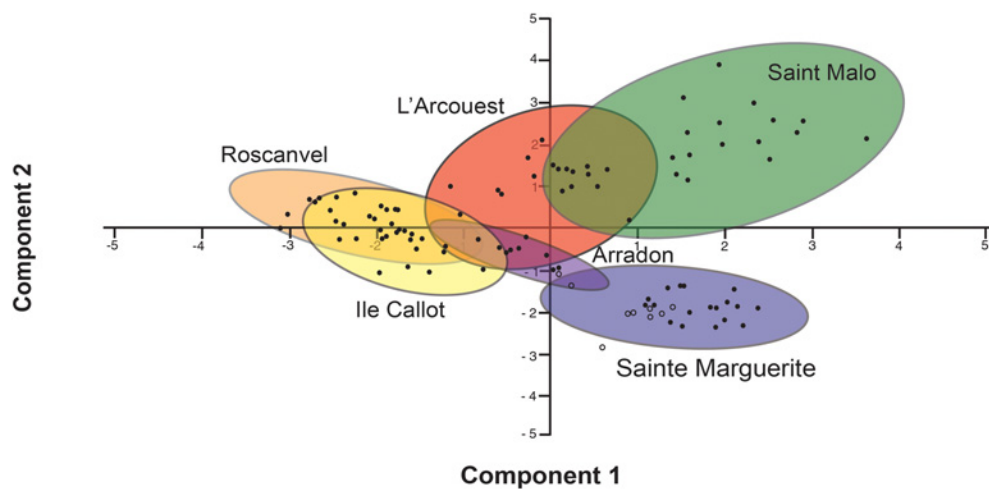
This study presents one of the first comparative analysis of inventories based on a dual morphological and molecular approaches for the same sampling sites, in addition to the inclusion of replicates to ensure statistically supported results. DNA metabarcoding of *Zostera marina* seagrass sediment uncovered vast taxonomic diversity, revealing that in a single snapshot, the invertebrate metazoan community diversity estimated by molecular surveys consisting of hundreds to thousands of MOTUs largely exceeded the 322 species uncovered by the morphological survey. In addition to this enhanced power of molecular barcode for unraveling the diversity of organisms present in the community, dissimilarities in the ranking of alpha diversities among meadows were also recorded (Fig. 6). This clearly illustrates the non-exhaustive nature of inventories obtained with both methods and their only partial overlap, confirming, despite the use of replicates, the extensive occurrence of false negative [56]. These differences, despite the comparison within the exact same locations, may stem from several reasons that are discussed more extensively below, among which the fact that each method requires samples of a different nature as a result of technical constraints. However, a congruent picture of community structure (beta diversities) was illustrated through spatial analyses, which identified statistically supported clusters corresponding to each meadow with both survey approaches, including the remote position of the highly differentiated Sainte Marguerite meadow community (Fig. 7).

The methods and results reported here provide another step towards the development of a powerful, repeatable and cost effective metabarcoding protocols for baseline studies integrated within environmental assessments while still pointing towards some required improvements. When compared to the morphological inventory, the observations of high diversity and the lack of some common species in the molecular inventory, as well as the amount of unknown MOTUs roughly assigned to related taxa, makes metabarcoding a powerful, but still partly blind method. Specifically, our difficulties to uncover a higher percentage of the most common species underlines i) the necessity of multi-gene surveys, ii) the further improvement of sampling protocols to allow the comparison of technically more similar samples between survey methods (rather than distinct sample collections from the exact same quadrates) and to include larger scale

A



B



**Fig 7. Principal Component Analyses (PCA) for the morphological and molecular surveys of meadows.** (A) PCA for the morphological inventory. (B) PCA for molecular inventory was performed using the COI gene. Each PCA was implemented using Jaccard dissimilarity matrices. Points are individual sediment samples taken at a specific core and mesh size, while colored ellipses are 95% confidence intervals for each meadow. For COI, the Sainte Marguerite cluster contains samples from 2010 (open circles) and 2011 (closed circles). Colors of the ellipses are specific to each meadow and these colors were maintained across both PCAs.

doi:10.1371/journal.pone.0117562.g007

spatial and temporal sampling, as well as iii) synergistic efforts between teams involved in molecular and morphological taxonomic identifications to feed, standardize and curate reference libraries. Additionally, the systematic molecular barcode of museum holotypes are already being performed at various museums around the world [57]. Awaiting the results of this global improvement, we advocate a double approach (the thorough comparison of morphological and

molecular inventories using “double described” species) at least during the first phase of biodiversity assessments, before safely engaging in longer-term ‘blind’ metabarcoding surveys.

## Advantages and disadvantages of both survey approaches

Advantages and disadvantages of both morphologic and molecular identification approaches have been discussed in detail recently [24,26,43]. A notable advantage of morphological screenings includes the ability to view whole specimens and to differentiate between biological footprints of transient organisms and those contemporarily present into the community. Further, the details provided by morphological screenings of organisms provide better taxonomic identifications for particular lineages. Morphological inventories also provide actual species abundance counts, whereas metabarcoding datasets are still limited to presence/absence data, given that technical factors likely influence the number of times a sequence is observed more than the abundance of the corresponding taxa [24].

The disadvantages of morphological approaches are amplified by dwindling taxonomic expertise and include difficulties with resolving cryptic species or meiofauna, as well as the complications of generating standardized protocols for assessment of broad community biodiversity. The time costs associated with taxonomic identifications has led to investigations for more time and cost efficient approaches for rapidly assessing ecosystem community structure [33]. Molecular methods, such as metabarcoding, have taken the forefront as solutions to the present dilemma, as these techniques have the ability to generate more extensive biological inventories while providing researchers with a standardized and economical approach to expeditiously assess community structure and biodiversity. Moreover, molecular collections taken with respect to organisms with body sizes less than 1mm can reveal species diversity that likely play a significant role in the functioning and stability of target communities, but is often difficult to retrieve via morphologic methods alone.

In the present study, comparisons of molecular and morphological surveys using only those samples obtained at  $\geq 1$ mm size fractions show molecular methods unmasking a higher amount of invertebrate diversity, providing a broader view of the taxa present in the community. We note that comparing only  $\geq 1$ mm size fractions provides a conservative estimate of diversity, however the finding of higher diversity is at least partially due to the enhanced ability of molecular methods to unmask the diversity of Annelida, which represent the dominant group of marine macrofauna [58]. Further, molecular surveys were particularly adept at uncovering animals common in smaller class sizes, such as nematodes (Fig. 5B), in addition to identifying “temporary meiofauna”—eggs or juveniles of larger animals found in the mesh size  $< 1$ mm [59] (Fig. 5A). Moreover, metabarcoding surveys may include the transitory presence of individuals having left biological footprints through partial remains such as scales, mucus, spines and feces. In some cases, mobile species native to the community might escape morphological surveys, as in the case of gastropods (such as those MOTUs found matching *Dendronotus frondosus*, *Goniodoris nodosa* and *Archidoris pseudoargus*), which were recorded in the molecular inventory but were not present in the morphological survey. It may be argued that molecular inventories could erroneously include incidental species (non-native transient diversity), however, the systematic inclusion of spatial replicates in the form of cores allows checking for the consistency of environmental collections.

## The synergistic relationship between molecular and morphological identifications

Molecular surveys (a combination of both genes) uncovered 36% common species in a single survey, suggesting that the other 64% were either i) not present in the sediment sampled at that

particular snapshot in time, ii) were present, but the tissue contained degraded DNA, iii) were not amplified with the primers employed, or iv) were not recognized through the assignment as adequately homologous to taxa present in reference databases. We also note that the low number of common species barcodes we obtained to add to the reference databases (14 for COI and 33 for 18S) could be due to mismatch in the primers or low quality of DNA from morphologically identified specimens that had been preserved part time in formaldehyde, which may have impaired the retrieval of common taxa among inventories.

The implementation of non-specific or “too-specific” primers could explain why the majority of common metazoans were not found. In the present study, we tested and implemented universal and degenerate primers for both COI and 18S, in order to increase our chances of retrieving the widest array of taxa possible. Despite our initial primer testing, the amplification of *numts*/pseudogenes and primer mismatching remain characteristics of the markers and technologies and we continue to work towards the improvement of these tools. Given the inherent characteristics of differing DNA types (evolutionary rate, neutrality, and presence of introns, for example), previous authors have suggested the use of a multilocus barcoding approach that includes mitochondrial and nuclear genes, to minimize gene/primer specific biases and enhance resolution at different taxonomic levels, a complementarity clearly depicted in [Fig. 3 \[15,60,61\]](#). Another reason for the lack of common species uncovered could be due to the analyses of only 10g of sediment taken from the sieved samples. Though our initial goal was to identify species present in a snapshot of sediment at a given time, replicate testing of sediment within each sample for consistency may determine if more species were present in the overall sample. However, in an analysis of earthworm communities, no differences were found between distinct DNA extractions performed on the same environmental sample [\[56\]](#).

Prior research described the problem of deriving accurate taxonomy using current databases, given the low reliability of publicly available taxonomic assignments past the class or family level, especially for organisms in smaller class sizes [\[6,25,26,43\]](#). As our sequence barcodes were taxonomically identified prior to being added to the reference database libraries, the inability to detect some of the common species if indeed present in the sediment, is not likely due to curation issues or misidentifications in the libraries used for assignments but more likely to technical pitfalls described above (the quality of DNA or primer matching). However, given metabarcoding’s reliance upon reference libraries for taxonomic assignments, our results also bring to light additional evidence for the absolute need for properly curated (i.e. taxonomically and molecularly identified) libraries to yield a higher likelihood of correct taxonomic assignment. This is the required condition to obtain inventories that are not only quantitatively superior for work on the spatial and temporal distribution of diversity, but also qualitatively reliable for understanding the taxonomic composition of communities and their evolution under differing environmental conditions.

We therefore suggest that pilot studies, including parallel morphological and molecular surveys, should be performed to i) formally test the efficiency of metabarcoding using the best protocols and genetic markers for the ecosystem studied, ii) provide a baseline for a temporal survey that could implement the blind metabarcoding method once taxonomic coverage and assignment are enhanced as well as the number of replicates determined for a reliable comparison, and iii) provide molecular databases against which future blind metabarcoding datasets could be compared, in order to survey the temporal evolution of a representative subset of the community. This would allow a baseline strategy with necessary parameters determined, prior to performing fully “blind” metabarcoding studies. Ultimately, refining the reference dataset to target specific groups would be useful in ecosystem investigations where the goals are to provide diversity estimates and community composition to the genus or species level based on reliable taxonomic inferences.

## Differences in relative alpha diversities

The discrepancy between the relative alpha diversities seen with the morphological and metabarcoding inventories across the different meadows is contrary to what has previously been observed [24]. This cannot be attributed to the inclusion of the meiofaunal compartment at constant “time costs” in the metabarcoding inventory, as similar results were obtained when including or excluding the smaller mesh size. Another explanation might be the nature of the collection methods and taxonomic indices upon which each approach relies. The morphological survey was performed using several collection devices and depended upon the presence of complete specimens, while the metabarcoding approach relied upon cores and typically identifies everything present in the environmental sample, including the remains of transient individuals. At first glance, the exposure of a community to transient individuals might prevent the accurate molecular assessment of true community structure, when visual surveys (i.e. morphological identifications) inherently filter transients and partial remains. However, metabarcoding assignments confirm that multiple cores result in similar community compositions within the six seagrass meadows, a finding that matches the morphological survey and supports the lack of spatial heterogeneity at the meadow scale. These findings suggest that each core can be considered a replicate for a representative description of communities.

Alpha diversity comparisons by meadow lacking strict correlation across surveys could also be the result of the variation in community composition in terms of phyla identified by molecular surveys (Figs. 4 and 5), which is magnified by the capacity of the molecular genes to uncover specific taxonomic groups. The 18S gene was particularly useful for identifying annelid polychaetes, a group which is known to dominate marine macrobenthos inventoried thus far [58], but is less thoroughly characterized by morphological surveys. The morphological survey showed high performance for identifying larger size class animals belonging to the phyla Chordata and Echinodermata, while underrepresenting phyla that can contain groups of smaller size classes (i.e. Annelida, Nematoda). Although this metabarcoding approach includes two types of markers to provide a more comprehensive estimates of biodiversity based on single surveys, the dominance of Annelida and Arthropoda provided by 18S and COI, respectively, likely had a large impact on the alpha diversities (Fig. 6).

The factors described above provide potential reasons for the observed differences in alpha diversity estimates of morphological versus molecular surveys, and the stronger power to uncover a broader range of taxa exhibited by the molecular surveys likely provide a more representative community inventory, despite a major disadvantage of not having species counts and thus quantitative data available.

## Comparison between COI and 18S barcoding markers

Results identified the 18S gene as uncovering higher numbers of common species as well as greater species diversity at the Sainte Marguerite and Arradon meadows, when compared to COI. We do not suggest that 18S is a unilaterally better marker, but in this case, had a better curated reference database and likely better primer matching owing to the composition of the communities studied, which were characterized by a high prevalence of taxa having better affinity to 18S primers. Further, we consider that the shorter fragment targeted by 18S (450bp versus 710bp with COI primers) may also explain a higher success when DNA quality is limited as with environmental samples. Both COI and 18S surveys still exhibited different and somehow complementary taxonomic specific biases, which pleads for multi-gene barcoding approaches. Future studies on different communities will continue to shed light on the intrinsic quality of each marker to unveil the taxonomic composition of a community as a whole, and

performing inventories using additional/new markers and primers will enhance complementarity with current markers and limit taxonomic bias.

## Community structure in morphological vs. molecular surveys

A similar pattern of clustering and meadow based community differentiation obtained with both survey methods was illustrated through multidimensional analysis (Fig. 7). While we do find some overlap of taxa within the seagrass communities, the stronger clustering of meadows shown by the molecular survey is likely the result of the weight given to rare taxa via the implementation of presence/absence analyses. Additionally, the molecular survey did identify higher levels of biodiversity than the morphological survey, and this more complete inventory provided a more holistic picture of community structure at *Z. marina* meadows. Furthermore, the analyses of 18S datasets from Sainte Marguerite and Arradon in 2010 also identify community differentiation amongst these two meadows (S2 Fig. and S3 Fig.). This discovery also provides confidence for temporal patterns and additional support for the reliability and repeatability of metabarcoding inventories for community analyses.

The remote position of the Sainte Marguerite communities was evident with both molecular and morphological approaches. The distinctiveness of the Sainte Marguerite community is likely due to more frequent disturbances in relation to its exposure to stronger oceanographic forces and patch extinctions. Sainte Marguerite is the least sheltered meadow that receives the highest wave energy as it is directly facing west, and its exposition to open sea currents may also explain a higher temporal variability. In 2008, local extinctions of large patches of this meadow were recorded, followed by progressive recolonization [35,62]. These demographic vagaries may be a result of oceanic influence and in turn impact the diversity and stability of associated communities. Interestingly, as sediment samples for this meadow were taken in 2010 and 2011, community composition appears to have changed only slightly over the course of a year, and more temporal studies may uncover whether the changes will continue with time elapsed since the last demographic crash.

## Conclusions

Results presented here support metabarcoding as an powerful and repeatable approach for expeditiously and cost effectively uncovering community structure and higher species richness than morphological methods, underlining the added value of the multi-gene approach to obtain a more balanced picture of represented taxa and overall community composition. The testing of a protocol that allowed replicate core sampling and analyses of specimens of distinct size classes to provide a more accurate view of community diversity illustrated that metabarcoding is consistent. Additionally, results raise substantial warnings as to the need for more direct comparisons, as well as comprehensive and rigorously curated reference libraries in order to shed light on the taxonomic identity of MOTUs detected through blind metabarcode. Results suggest that a first step consisting of “double inventories” (morphological and molecular) using standardized protocols may be required, in order to check the accuracy of molecular protocols for the targeted communities, in addition to supplying a reference inventory and molecular database for subsequent “blind metabarcode” temporal surveys.

## Supporting Information

**S1 Dataset. Morphological survey of 322 species found at six *Zostera marina* seagrass meadows in 2011, identified by IUEM/REBENT.** Each meadow contains collections from three cores, and presence of species is indicated by “1”, while absence is indicated by “0”. The most common species found in this inventory are highlighted in yellow, while those common

species *not* found in the inventory are highlighted in light green. Columns for those species uncovered by each gene are highlighted in dark green for 18S and dark blue for COI.

(XLSX)

**S2 Dataset. Global BLAST results for 12,548 unassigned COI MOTUs.** Included are MOTU names, the GenBank identifying number (GI), the e-value, the closest species match, the kingdom, the phylum (for Eukarya only), and the frequencies of each kingdom present in the dataset. The blastn algorithm was implemented under default parameters (see Appendix C of the BLAST command line applications user manual, <http://www.ncbi.nlm.nih.gov/books/NBK1763/>).

(XLSX)

**S3 Dataset. Global BLAST results for 48 unassigned 18S MOTUs.** Included are MOTU names, the GenBank identifying number (GI), the e-value, the closest species match, the kingdom, the phylum, and the frequencies of each kingdom present in the dataset. The blastn algorithm was implemented under default parameters (see Appendix C of the BLAST command line applications user manual, <http://www.ncbi.nlm.nih.gov/books/NBK1763/>).

(XLSX)

**S4 Dataset. Frequencies of taxa identified by metazoan phyla by each respective survey method.**

(XLSX)

**S5 Dataset. Chao1 (alpha diversity) and Jaccard distances (beta diversity) for morphological survey inventoried by REBENT.** Chao1 values were generated by meadow site using Vegan in the software R, while Jaccard distances were generated by meadow, by core and by sample using the program PAST.

(XLSX)

**S6 Dataset. Chao1 (alpha diversity) and Jaccard distances (beta diversity) for the COI gene molecular inventory.** Chao1 values were generated by meadow site using Vegan in the software R, while Jaccard distances were generated by meadow, by core and by sample using the program PAST.

(XLSX)

**S7 Dataset. Chao1 (alpha diversity) and Jaccard distances (beta diversity) for the 18S gene molecular inventory.** Chao1 values were generated by meadow site using Vegan in the software R, while Jaccard distances were generated by meadow, by core and by sample using the program PAST.

(XLSX)

**S1 Fig. (unassigned MOTUs included) Principal Components Analysis (PCA) using the Jaccard coefficient, for six meadows described with the COI gene.**

(PDF)

**S2 Fig. (unassigned MOTUs included) Principal Components Analysis (PCA) using the Jaccard coefficient, for Sainte Marguerite and Arradon meadows described with the 18S gene.**

(PDF)

**S3 Fig. (unassigned MOTUs excluded) Principal Components Analysis (PCA) using the Jaccard coefficient, for Sainte Marguerite and Arradon meadows described with the 18S gene.**

(PDF)

**S1 Protocol. DNA extraction and sequencing protocols for REBENT identified metazoans.**  
(DOCX)

**S2 Protocol. QIIME scripts used for bioinformatics data processing analyses.**  
(DOCX)

**S1 Table. Characteristics of Roche 454 FLX produced amplicons before and after filtering stages.**  
(DOCX)

**S2 Table. (unassigned MOTUs included) One and two-way ANOSIM results implemented in PAST under the Jaccard similarity index for 18S and COI under  $n = 9999$  permutations.**  
(DOCX)

**S3 Table. (unassigned MOTUs excluded) One and two-way ANOSIM results implemented in PAST under the Jaccard similarity index for 18S and COI under  $n = 9999$  permutations.**  
(DOCX)

**S4 Table. (unassigned MOTUs included) One-way ANOSIM pairwise matrix p-values for meadow and mesh size in COI. SMG = Sainte Marguerite.**  
(DOCX)

**S5 Table. (unassigned MOTUs excluded) One-way ANOSIM pairwise matrix p-values for meadow and mesh size in COI. SMG = Sainte Marguerite.**  
(DOCX)

**S6 Table. (unassigned MOTUs included) One-way ANOSIM pairwise matrix p-values for meadow and mesh size in 18S**  
(DOCX)

**S7 Table. (unassigned MOTUs excluded) One-way ANOSIM pairwise matrix p-values for meadow and mesh size in 18S**  
(DOCX)

**S8 Table. Mantel test results implemented in PAST for morphology and molecular (COI) Jaccard distances by meadow, under the Bray-Curtis index and  $n = 9999$  permutations. Results with unassigned COI MOTUs included and excluded are reported**  
(DOCX)

**S9 Table. One and two-way ANOSIM results implemented in PAST under the Jaccard similarity index for the morphology dataset under  $n = 9999$  permutations.**  
(DOCX)

**S10 Table. One-way ANOSIM pairwise matrix p-values for meadows in the morphology dataset. SMG = Sainte Marguerite.**  
(DOCX)

## Acknowledgments

We would like to extend our thanks to Olivier Soubigou, Julien Chopelet and Christian Hily for technical help and useful discussions throughout the course of this project. We also thank Lenaick Menot and Ronan Becheler for their contributions to the field sampling for the study. Finally, we would like to thank the Rebent-Bretagne and particularly the Region Bretagne, the European Regional Development Fund, Ifremer, and the Agence de l'Eau Loire Bretagne for funding the morphological identifications survey.



## Author Contributions

Conceived and designed the experiments: DAC JM SAH. Performed the experiments: SAH OM JG MM. Analyzed the data: DAC MP SAH. Contributed reagents/materials/analysis tools: MP SAH. Wrote the paper: DAC SAH. Performed field and molecular work: SAH OM. Performed collections and analyses of morphological datasets: JG MM. Contributed to the final manuscript: DAC MP OM MM JG JM SAH.

## References

1. Pimm S, Russell GJ, Gittleman JL, Brooks TM (1995) The future of biodiversity. *Science* 269: 347–350. PMID: [17841251](#)
2. Barnosky A, Matzke N, Tomiya A, Wogan GOU, Swartz B, et al. (2011) Has the Earth's sixth mass extinction already arrived? *Nature* 471: 51–57. doi: [10.1038/nature09678](#) PMID: [21368823](#)
3. Pereira H, Ferrier S, Walters M, Geller GN, Jongman RHG, et al. (2013) Essential biodiversity variables. *Science* 339: 277–278. doi: [10.1126/science.1229931](#) PMID: [23329036](#)
4. Bucklin A, Steinke D, Blanco-Bercial L (2011) DNA barcoding of marine metazoa. *Annual Review of Marine Science* 3: 471–508. PMID: [21329214](#)
5. Tang C, Leasi F, Obertegger U, Kieneke A, Barraclough TG et al. (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: 16208–16212. doi: [10.1073/pnas.1209160109](#) PMID: [22988084](#)
6. Lindeque P, Parry HE, Harmer RA, Somerfield PJ, Atkinson A (2013) Next Generation Sequencing Reveals the Hidden Diversity of Zooplankton Assemblages. *PLoS One* 8: e81327. doi: [10.1371/journal.pone.0081327](#) PMID: [24244737](#)
7. Pawlowski J, Christen R, Lecroq B, Bachar D, Shahbazkia HR, et al. (2011) Eukaryotic richness in the abyss: insights from pyrotag sequencing. *PLoS One* 6: e18169. doi: [10.1371/journal.pone.0018169](#) PMID: [21483744](#)
8. Scheffers B, Joppa LN, Pimm SL, Laurance WF (2012) What we know and don't know about Earth's missing biodiversity. *Trends in ecology & evolution* 27: 501–510. doi: [10.1038/ncomms7092](#) PMID: [25600177](#)
9. Hebert P, Cywinska A, Ball SL (2003b) Biological identifications through DNA barcodes. *Proc R Soc Lond B Biol Sci* 270: 313–321.
10. Puillandre N, Lambert A, Brouillet S, Achaz G (2012) ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Molecular Ecology* 21: 1864–1877. doi: [10.1111/j.1365-294X.2011.05239.x](#) PMID: [21883587](#)
11. Blaxter M, Mann J, Chapman T, Thomas F, Whitton C, et al. (2005) Defining operational taxonomic units using DNA barcode data. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360: 1935–1943. PMID: [16214751](#)
12. Markmann M, Tautz D (2005) Reverse taxonomy: an approach towards determining the diversity of meiobenthic organisms based on ribosomal RNA signature sequences. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360: 1917–1924. PMID: [16214749](#)
13. Newmaster S, Fazekas AJ, Ragupathy S (2006) DNA barcoding in land plants: evaluation of rbcL in a multigene tiered approach. *Botany* 84: 335–341.
14. Miller S (2007) DNA barcoding and the renaissance of taxonomy. *Proceedings of the National Academy of Sciences* 104: 4775–4776. PMID: [17363473](#)
15. Frézal L, Leblois R (2008) 4 years of DNA barcoding: current advances and prospects. *Infection, Genetics and Evolution* 8: 727–736. doi: [10.1016/j.meegid.2008.05.005](#) PMID: [18573351](#)
16. Hajibabaei M, Shokralla S, Zhou X, Singer GA, Baird DJ (2011) Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. *PLoS One* 6: e17497. doi: [10.1371/journal.pone.0017497](#) PMID: [21533287](#)
17. Travis J, Coleman FC, Auster PJ, Cury PM, Estes JM, et al. (2014) Integrating the invisible fabric of nature into fisheries management. *Proceedings of the National Academy of Sciences* 111: 581–584. doi: [10.1073/pnas.1305853111](#) PMID: [24367087](#)
18. Veraart A, Faassen EJ, Dakos V, van Nes EH, Lurling M, et al. (2012) Recovery rates reflect distance to a tipping point in a living system. *Nature* 481: 357–359. doi: [10.1038/nature10723](#) PMID: [22198671](#)
19. Jain S, Krishna S (2001) A model for the emergence of cooperation, interdependence, and structure in evolving networks. *Proceedings of the National Academy of Sciences* 98: 543–547.

20. Whitham T, Young WP, Martinsen GD, Gehring CA, Schweitzer JA, et al. (2003) Community ecosystem genetics: a consequence of the extended phenotype. *Ecology* 84: 559–573.
21. Baird D, Hajibabaei M (2012) Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. *Molecular Ecology* 21: 2039–2044. PMID: [22590728](#)
22. Yu D, Ji Y, Emerson BC, Wang X, Ye C, et al. (2012) Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution* 3: 613–623.
23. Barber P, Boyce SL (2006) Estimating diversity of Indo-Pacific coral reef stomatopods through DNA barcoding of stomatopod larvae. *Proceedings of the Royal Society B: Biological Sciences* 273: 2053–2061. PMID: [16846913](#)
24. Ji Y, Ashton L, Pedley SM, Edwards DP, Tang Y, et al. (2013) Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology letters* 16: 1245–1257. doi: [10.1111/ele.12162](#) PMID: [23910579](#)
25. Creer S, Fonseca VG, Porazinska DL, Giblin-Davis RM, Sung W, et al. (2010) Ultrasequencing of the meiofaunal biosphere: practice, pitfalls and promises. *Molecular Ecology* 19: 4–20. doi: [10.1111/j.1365-294X.2009.04473.x](#) PMID: [20331766](#)
26. Fonseca V, Carvalho GR, Sung W, Johnson HF, Power DM, et al. (2010) Second-generation environmental sequencing unmasks marine metazoan biodiversity. *Nature Communications* 1: 98. doi: [10.1038/ncomms1095](#) PMID: [20981026](#)
27. Thomsen P, Kielgast J, Iversen LL, Møller PR, Rasmussen M, et al. (2012) Detection of a diverse marine fish fauna using environmental DNA from seawater samples. *PLoS One* 7: e41732. doi: [10.1371/journal.pone.0041732](#) PMID: [22952584](#)
28. Porazinska D, Giblin-Davis RM, Faller L, Farmerie W, Kanzaki N, et al. (2009) Evaluating high-throughput sequencing as a method for metagenomic analysis of nematode diversity. *Molecular Ecology Resources* 9: 1439–1450. doi: [10.1111/j.1755-0998.2009.02611.x](#) PMID: [21564930](#)
29. Porazinska D, Giblin-Davis RM, Esquivel A, Powers TO, Sung W, et al. (2010) Ecometagenetics confirm high tropical rainforest nematode diversity. *Molecular Ecology* 19: 5521–5530. doi: [10.1111/j.1365-294X.2010.04891.x](#) PMID: [21054606](#)
30. Valentini A, Pompanon F, Taberlet P (2009) DNA barcoding for ecologists. *Trends in Ecology & Evolution* 24: 110–117. doi: [10.1038/ncomms7092](#) PMID: [25600177](#)
31. Leray M, Boehm JT, Mills SC, Meyer CP (2012) Moorea BIOCDE barcode library as a tool for understanding predator–prey interactions: insights into the diet of common predatory coral reef fishes. *Coral Reefs* 31: 383–388.
32. Stoeck T, Bass D, Nebel M, Christen R, Jones MD, et al. (2010) Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Molecular Ecology* 19: 21–31. doi: [10.1111/j.1365-294X.2009.04480.x](#) PMID: [20331767](#)
33. Taberlet P, Coissac E, Hajibabaei M, Rieseberg LH (2012) Environmental DNA. *Molecular Ecology* 21: 1789–1793. doi: [10.1111/j.1365-294X.2012.05542.x](#) PMID: [22486819](#)
34. REBENT (2010) REseau BENThique. [www.rebent.org](http://www.rebent.org)
35. Becheler R, Diekmann O, Hily C, Moalic Y, Arnaud-Haond S (2010) The concept of population in clonal organisms: mosaics of temporally colonized patches are forming highly diverse meadows of *Zostera marina* in Brittany. *Molecular Ecology* 19: 2394–2407. doi: [10.1111/j.1365-294X.2010.04649.x](#) PMID: [20465589](#)
36. 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: 294–299. PMID: [7881515](#)
37. Blaxter M, De Ley P, Garey JR, Liu LX, Scheldeman P, et al. (1998) A molecular evolutionary framework for the phylum Nematoda. *Nature* 392: 71–75. PMID: [9510248](#)
38. Meyer C (2003) Molecular systematics of cowries (Gastropoda: Cypraeidae) and diversification patterns in the tropics. *Biological Journal of the Linnean Society* 79: 401–459.
39. Hebert PDN, Ratnasingham S, de Waard JR (2003a) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London Part B* 270: S96–S99.
40. Caporaso J, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* 7: 335–336. doi: [10.1038/nmeth.f.303](#) PMID: [20383131](#)
41. Floyd R, Abebe E, Papert A, Blaxter M (2002) Molecular barcodes for soil nematode identification. *Molecular Ecology* 11: 839–850. PMID: [11972769](#)

42. Edgar R (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26: 2460–2461. doi: [10.1093/bioinformatics/btq461](https://doi.org/10.1093/bioinformatics/btq461) PMID: [20709691](https://pubmed.ncbi.nlm.nih.gov/20709691/)
43. Bik H, Porazinska DL, Creer S, Caporaso JG, Knight R, et al. (2012) Sequencing our way towards understanding global eukaryotic biodiversity. *Trends in ecology & evolution* 27: 233–243. doi: [10.1038/ncomms7092](https://doi.org/10.1038/ncomms7092) PMID: [25600177](https://pubmed.ncbi.nlm.nih.gov/25600177/)
44. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, et al. (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic acids research* 35: 7188–7196. PMID: [17947321](https://pubmed.ncbi.nlm.nih.gov/17947321/)
45. Altschul S, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology* 215: 403–310. PMID: [2231712](https://pubmed.ncbi.nlm.nih.gov/2231712/)
46. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, et al. (2009) BLAST+: architecture and applications. *BMC bioinformatics* 10: 421. doi: [10.1186/1471-2105-10-421](https://doi.org/10.1186/1471-2105-10-421) PMID: [20003500](https://pubmed.ncbi.nlm.nih.gov/20003500/)
47. Uhlig G, Thiel H, Gray JS (1973) The quantitative separation of meiofauna. *Helgoländer wissenschaftliche Meeresuntersuchungen* 25: 173–195. doi: [10.1111/1467-9566.12160](https://doi.org/10.1111/1467-9566.12160) PMID: [25470322](https://pubmed.ncbi.nlm.nih.gov/25470322/)
48. Chao A (1984) Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics* 11: 265–270.
49. Hughes J, Hellman JJ, Ricketts TH, Bohannon BJM (2001) Counting the uncountable: statistical approaches to estimating microbial diversity. *Applied and Environmental Microbiology* 67: 4399–4406. PMID: [11571135](https://pubmed.ncbi.nlm.nih.gov/11571135/)
50. R Development Core Team (2011) R: A Language and Environment for Statistical Computing. In: *Computing TRFfS*, editor. Vienna, Austria. doi: [10.1016/j.neuroimage.2011.01.013](https://doi.org/10.1016/j.neuroimage.2011.01.013) PMID: [21238596](https://pubmed.ncbi.nlm.nih.gov/21238596/)
51. Hammer Ø, Harper DAT, Ryan PD (2001) PAST: Paleontological Statistics Software Package for Education and Data Analysis *Palaeontol. Electronica* 4: 1–9.
52. Letunic I, Bork P (2007) Interactive Tree Of Life (iTOL): An online tool for phylogenetic tree display and annotation. *Bioinformatics* 23: 127–128. PMID: [17050570](https://pubmed.ncbi.nlm.nih.gov/17050570/)
53. Berney C, Fahmi J, Pawlowski J (2004) How many novel eukaryotic 'kingdoms'? Pitfalls and limitations of environmental DNA surveys. *BMC Biology* 2: 1–13. PMID: [14731304](https://pubmed.ncbi.nlm.nih.gov/14731304/)
54. Buhay J (2009) "COI-like" sequences are becoming problematic in molecular systematic and DNA barcoding studies. *Journal of Crustacean Biology* 29: 96–110.
55. Bik H, Sung W, De Ley P, Baldwin JG, Sharma J, et al. (2011) Metagenetic community analysis of microbial eukaryotes illuminates biogeographic patterns in deep-sea and shallow water sediments. *Molecular Ecology* 21: 1048–1059. doi: [10.1111/j.1365-294X.2011.05297.x](https://doi.org/10.1111/j.1365-294X.2011.05297.x) PMID: [21985648](https://pubmed.ncbi.nlm.nih.gov/21985648/)
56. Ficetola GF, Pansu J, Bonin A, Coissac E, Giguet-Covex C, et al. (2014) Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Molecular Ecology Resources*: n/a-n/a. doi: [10.1111/1755-0998.12366](https://doi.org/10.1111/1755-0998.12366) PMID: [25545675](https://pubmed.ncbi.nlm.nih.gov/25545675/)
57. Puillandre N, Bouchet P, Boisselier-Dubayle MC, Brisset J, Buge B, et al. (2012) New taxonomy and old collections: integrating DNA barcoding into collections curation processes. *Molecular Ecology Resources* 12. doi: [10.1111/1755-0998.12018](https://doi.org/10.1111/1755-0998.12018) PMID: [23227485](https://pubmed.ncbi.nlm.nih.gov/23227485/)
58. Struck T, Paul C, Hill N, Hartmann S, Hosel C, et al. (2011) Phylogenomic analyses unravel annelid evolution. *Nature* 471: 95–98. doi: [10.1038/nature09864](https://doi.org/10.1038/nature09864) PMID: [21368831](https://pubmed.ncbi.nlm.nih.gov/21368831/)
59. McIntyre AD (1969) Ecology of marine meiobenthos. *Biological Reviews* 44: 245–288.
60. Hajibabaei M, Singer GAC, Hebert PDN, Hickey DA (2007) DNA barcoding: how it complements taxonomy, molecular phylogenetics and population genetics. *TRENDS in Genetics* 23: 167–172. PMID: [17316886](https://pubmed.ncbi.nlm.nih.gov/17316886/)
61. Smith M, Poyarkov NA, Hebert PDN (2008) CO1 DNA barcoding amphibians: take the chance, meet the challenge. *Molecular Ecology Resources* 8: 235–246. doi: [10.1111/j.1471-8286.2007.01964.x](https://doi.org/10.1111/j.1471-8286.2007.01964.x) PMID: [21585765](https://pubmed.ncbi.nlm.nih.gov/21585765/)
62. Becheler R, Benkara E, Moalic Y, Hily C, Arnaud-Haond S (2014) Scaling of processes shaping the clonal dynamics and genetic mosaic of seagrasses through temporal genetic monitoring. *Heredity* 112: 114–121. doi: [10.1038/hdy.2013.82](https://doi.org/10.1038/hdy.2013.82) PMID: [24022498](https://pubmed.ncbi.nlm.nih.gov/24022498/)