University of Southampton Research Repository

Copyright © and Moral Rights for this thesis and, where applicable, any accompanying data are retained by the author and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This thesis and the accompanying data cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder/s. The content of the thesis and accompanying research data (where applicable) must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holder/s.

When referring to this thesis and any accompanying data, full bibliographic details must be given, e.g.

Thesis: Author (Year of Submission) "Full thesis title", University of Southampton, name of the University Faculty or School or Department, PhD Thesis, pagination.

Data: Author (Year) Title. URI [dataset]

University of Southampton

Faculty of Environmental and Life Sciences

School of Ocean and Earth Science

Environmental DNA metabarcoding and the study of marine biodiversity and biological invasions

Luke Earl Holman

ORCID ID 0000-0002-8139-3760

Thesis for the degree of Doctor of Philosophy

April 2021

University of Southampton

Abstract

Faculty of Faculty of Environmental and Life Sciences
School of Ocean and Earth Sciences
Doctor of Philosophy
Environmental DNA metabarcoding and the study of marine biodiversity and biological invasions
by
Luke Earl Holman

Human activities negatively impact the biodiversity of oceanic and coastal ecosystems. Studies show that marine biodiversity is linked to ecosystem function and in turn to the ecosystem services upon which humans depend. A key driver of biodiversity loss is the introduction of non-native species which have a dramatic negative effect on biodiversity and carry a vast global economic cost. It is therefore critical to understand the causes of anthropogenic biodiversity loss, monitor the effect of management actions and, ultimately, document ecosystem recovery. The use of DNA isolated from enviromental samples such as water or sediment (environmental DNA or eDNA) to infer the presence or absence of an organism is revolutionising our ability to monitor aquatic ecosystems. Yet, we have a limited understanding how eDNA can be deployed to answer fundamental and applied questions in marine biodiversity and invasion biology.

This thesis begins by showing eDNA metabarcoding to be an accurate and sensitive tool for detection of marine non-native species and demonstrates how methodological choices affect the biodiversity that is detected. The utility of eDNA metabarcoding in broad-scale biogeographic analyses is then shown by identifying that marine animals, protists and bacteria can have remarkably similar biogeographic patterns. The thesis then reveals the limitations of eDNA metabarcoding data in comparison to manual biodiversity surveys and mitochondrial DNA sequences isolated from tissue samples for several non-native species. Finally, it develops a method for non-invasively genotyping organisms using eDNA, showing that eDNA contains both mitochondrial and nuclear genetic biodiversity.

Overall, this thesis leverages the sensitivity, accuracy and taxonomic breadth of eDNA-based tools to answer key questions in invasion biology and marine biodiversity.

Table of Contents

Table of Contents .... Error! Bookmark not defined.

Research Thesis: Declaration of Authorship v

Statement of Contribution vii

Acknowledgements ix

Definitions and Abbreviations xi

Chapter 1 Introduction 1

1.1 Marine biodiversity 1

1.1.1 What is biodiversity? 1
1.1.2 Biogeographic regions & ecosystem services 3
1.1.3 Threats to biodiversity 5

1.2 Invasive species 7

1.2.1 Defining the invader 7
1.2.2 Invasion process 9
1.2.3 Marine vectors 10
1.2.4 Marine impacts 11
1.2.5 Current issues & opportunities 13

1.3 Environmental DNA 14

1.3.1 What is eDNA? 14
1.3.2 The ecology of eDNA 15

1.3.2.1 Origin 16
1.3.2.2 State 18
1.3.2.3 Transport 18
1.3.2.4 Fate 19
1.3.2.5 Limitations of the ecology of eDNA 21

1.3.3 eDNA detection methods 22

1.3.3.1 Single-species targeted detection 22
1.3.3.2 Multi-species passive detection 23

1.3.4 eDNA in the oceans 25

1.4 Thesis aims & overview 27
1.5 References 29

Chapter 2 Detection of introduced and resident marine species using environmental DNA metabarcoding of sediment and water 41

2.1 Citation 41
2.2 Published Article 42
2.3 Supplementary Information 52
  2.3.1 Supplementary Information 1 52
  2.3.2 Supplementary Information 2 53
  2.3.3 Supplementary Information 3 55
  2.3.4 Supplementary Information 4: 56
  2.3.5 Supplementary Information 5 57
  2.3.6 Supplementary Information 6 58

Chapter 3 Animals, protists and bacteria share marine biogeographic patterns 59

3.1 Citation 59
3.2 Abstract 60
3.3 Introduction 61
3.4 Results 64
  3.4.1 DNA sequencing 64
  3.4.2 Taxonomic assignments & alpha diversity 64
  3.4.3 Beta diversity 67
  3.4.4 Distance-decay 70
3.5 Discussion 72
3.6 Methods 76
  3.6.1 Field sampling 76
  3.6.2 Environmental DNA extraction 76
  3.6.3 High throughput eDNA amplicon sequencing 77
  3.6.4 Bioinformatics 78
  3.6.5 Environmental, human impact and geographic data 80
  3.6.6 Ecological statistics 81
3.7 References 84
3.8 Supplementary Information 88
  3.8.1 Supplementary Figure 1 88
  3.8.2 Supplementary Figure 2 89
  3.8.3 Supplementary Table 1 90
  3.8.4 Supplementary Table 2 91
  3.8.5 Supplementary Table 3 93
  3.8.6 Supplementary Table 4 95
  3.8.7 Supplementary Table 5 97

3.8.8 Supplementary Table 6 99
3.8.9 Supplementary Table 7 101
3.8.10 Supplementary Table 8 103
3.8.11 Supplementary Table 9 105
3.8.12 Supplementary Note 1 106
3.8.13 Supplementary Note 2 108
3.8.14 Supplementary Note 3 112
3.8.15 Supplementary Note 4 116
3.8.16 Supplementary Note 5 122
3.8.17 Supplementary References 125

Chapter 4 Managing human mediated range shifts: understanding spatial temporal and genetic variation in marine non-native species.....127

4.1 Citation 127
4.2 Abstract 128
4.3 Introduction 129
4.4 Methods 132
  4.4.1 Fieldwork and historical biodiversity data 132
  4.4.2 Sample collection, DNA extraction and sequencing 133
  4.4.3 Environmental DNA metabarcoding 134
  4.4.4 Data manipulation and statistical analysis 136
4.5 Results 137
  4.5.1 Range shifts 137
  4.5.2 Changes in genetic composition 138
4.6 Discussion 142
4.7 References 146
4.8 Supplementary Information 151
  4.8.1 Supplementary Table 1 151
  4.8.2 Supplementary Table 2 152
  4.8.3 Supplementary Table 3 153
  4.8.4 Supplementary Note 1 154
  4.8.5 Supplementary Note 2 156
  4.8.6 Supplementary Note 3 158
  4.8.7 Supplementary References 160

Chapter 5 Demonstration of the use of environmental DNA for the non-invasive genotyping of a bivalve mollusk, the European Flat Oyster (*Ostrea edulis*) 161

5.1 Citation 161
5.2 Supplementary Information 169
5.2.1 Supplementary Figure 1 169
5.2.2 Supplementary Figure 2 170

Chapter 6 Synthesis & Conclusions 171

6.1 Overview 171
6.2 Synthesis 174
6.2.1 A revolution in marine ecosystem monitoring 174
6.2.2 Non-native species and eDNA 175
6.2.3 Primer choice and eDNA metabarcoding 177
6.3 Future Work 179
6.4 References 182

Research Thesis: Declaration of Authorship

Print name: Luke Earl Holman

Title of thesis: Environmental DNA metabarcoding and the study of marine biodiversity and biological invasions

I declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself below;
7. Parts of this work have been published as:

Holman LE, de Bruyn M, Creer S, Carvalho G, Robidart, J, Rius M (2019) The detection of novel and resident marine non-indigenous species using environmental DNA metabarcoding of seawater and sediment. *Scientific Reports* 9:11559

Holman LE, de Bruyn M, Creer S, Carvalho G, Robidart, J, Rius M (2021) Animals, protists and bacteria share marine biogeographic patterns. *Nature Ecology and Evolution* 5 (6), 738–746

Holman LE, Hollenbeck CM, Ashton TJ, Johnston IA (2019*)* Demonstration of the Use of Environmental DNA for the Non-Invasive Genotyping of a Bivalve Mollusk, the European Flat Oyster (Ostrea edulis). *Frontiers in Genetics*. 10:1159

Signature:.. ....Date:

Statement of Contribution

This thesis contains contributions from co-authors and collaborators as detailed below for each chapter.

Chapter One: Luke Holman, Marc Rius
LH wrote the text and produced all unpublished diagrams. MR provided suggestions on the initial draft.

Chapter Two: Luke Holman, Mark de Bruyn, Simon Creer, Gary Carvalho, Julie Robidart, Marc Rius
LH and MR designed the experiment with input from all authors. LH collected the samples, performed the lab work, analysed the data, prepared all figures and produced a first draft of the manuscript. All authors provided guidance on the interpretation and analysis of the data. All authors critically contributed to the final manuscript.

Chapter Three: Luke Holman, Mark de Bruyn, Simon Creer, Gary Carvalho, Julie Robidart, Marc Rius
LH and MR designed the study. LH collected the samples, performed the laboratory work, analysed the data, prepared all figures and wrote the first draft of the manuscript. All authors provided guidance on the interpretation and analysis of the data. All authors contributed to subsequent drafts of the manuscript.

Chapter Four: Luke Holman, Shirley Parker-Nance, Mark de Bruyn, Simon Creer, Gary Carvalho, Marc Rius
LH and MR designed the study. LH collected the samples, performed the lab work, analysed the data, prepared all figures and wrote the first draft of the manuscript. SPN provided in country assistance with fieldwork. MR provided guidance on the interpretation and analysis of the data. All authors critically contributed to the final manuscript.

Chapter Five: Luke Holman, Christopher Hollenbeck, Tom Ashton, Ian Johnston

LH designed the experiment with input from TA and IJ. LH performed the lab work and data analysis. LH and CH prepared a first draft of the manuscript, all authors critically contributed to subsequent drafts of the manuscript.

Chapter Six: Luke Holman

LH is the sole author.

Acknowledgements

I'd first like to thank my supervisor Marc Rius who has always made time to work on a manuscript, offer an opinion or just chat about science over lunch. He has generously offered support, guidance and mentorship while also giving me the freedom to develop as an independent scientist. Thank you Marc.

I will miss the monthly calls with my co-supervisors Mark de Bruyn, Si Creer and Gary Carvalho. Thank you all for your contributions to this work, for meeting with me frequently and for your support and advice.

I'm grateful to Julie Robidart for joining my supervisory team along the way, thank you for igniting an interest in microbes I didn't know I had, and for welcoming me into the lab.

Research in South Africa would not have been possible without Shirley Parker-Nance who was generous, kind and shared her knowledge of South African ascidians. Thank you. I'm also grateful to Molly Czachur and Thomas Grevesse for sharing their enthusiasm and energy during long days in the field.

Thanks go to Ian Johnston and Tom Ashton at Xelect Ltd. for supporting me through a CASE studentship and for their contributions to our manuscript. I am a better scientist and writer for having worked with them. Thanks also to the whole team at Xelect and in particular Chris Hollenbeck for sharing cheese toasties and interesting science with me.

Aubrie and Joe Onoufriou always made me feel at home when up in Fife, thank you both.

Warm thanks to the many friends I met along the way for their companionship and encouragement: the past and present members of the Ecology and Evolution Lab, the Ocean Technology and Engineering Group and Café 566/13. Particular thanks go to Jamie Hudson for always finding fun in the technical and for sharing in my enthusiasm for the aesthetics of science, Christina Wood and Nic Pratt for New Forest weekends and pub lunches, Zoë Holbrook for always making time for a coffee break and Jake Stone for beers on sunny evenings.

Thank you to my family and especially my parents Cathy and Chris Holman for their support, I could not have made it this far without them.

Final thanks to Charlie Davis who has been by my side for every high and low. Our weekends in wild places permeate my memories of this work.

Thank you all.

Definitions and Abbreviations

16S	**16S** small subunit ribosomal RNA (prokaryotic)
18S	**18S** small subunit ribosomal RNA (eukaryotic)
ANOVA	**an**alysis **o**f **va**riance
BCE	**b**efore the **c**ommon **e**ra
bp	nucleotide **b**ase **p**airs
CAS	CRISPR **as**sociated protein
COI	**c**ytochrome c **o**xidase subunit **I**
CRISPR	**c**lustered **r**egularly **i**nterspaced **s**hort **p**alindromic **r**epeats
dbRDA	**d**istance-**b**ased **r**e**d**undancy **a**nalysis
DNA	**d**eoxyribo**n**ucleic **a**cid
eDNA	**e**nvironmental **DNA**
ddPCR	**d**igital **d**roplet **PCR**
NCBI	**N**ational **C**enter for **B**iotechnology **I**nformation (US)
NNS	**n**on-**n**ative **s**pecies
PCR	**p**olymerase **c**hain **r**eaction
PERMANOVA	**per**mutational **m**ultivariate **an**alysis **o**f **va**riance
qPCR	**q**uantitative **PCR**
SSS	**s**ea **s**urface **s**alinity
SST	**s**ea **s**urface **t**emperature

Chapter 1 Introduction

1.1 Marine biodiversity

1.1.1 What is biodiversity?

One of the most remarkable features of life on earth is its diversity. As we stare into a rockpool, under an upturned stone or down the microscope we see a world characterised by great variation in forms. The study of biological variation has been a feature of scientific endeavour for millennia. In his *History of Animals* (Latin: Historia Animalium), published in the 4th Century BCE, Aristotle recorded the great variety of life. Notably, he attempted to delineate his observations according to characteristics, such as animals with or without red blood. Arguably the next great advance in our understanding of biological diversity was the taxonomy of Carl Linnaeus who sought to impose order on the natural world with a system for classifying life presented in *Systema Naturae* (although only plants were considered in this publication). This method, now known as Linnaean taxonomy, provided a framework for subsequent research, and enabled early naturalists to communicate their discoveries and descriptions of life.

Although our understanding of the variety and volume of life increased rapidly over recent centuries the term *Biological Diversity* is new. The earliest recognised (Magurran 2013) use of the term can be found in Gerbilskii and Petrunkevitch (1955). The word Biodiversity is attributed to Walter G. Rosen (Wilson 1988) (then written as *BioDiversity)* at the eponymous 'National Forum of BioDiversity' held in Washington DC (USA) in 1986. Use of the term subsequently increased, both in academic discourse and in the minds, and legislation, of policy makers. Indeed, the most commonly used definition of biodiversity comes from the landmark United Nations Convention on Biological Diversity (1992);

> ' "Biological diversity" means the variability among living organisms from all sources including, inter alia, terrestrial, marine and other aquatic ecosystems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems. ' *Pg. 3 Convention on Biological Diversity (1992)*

This definition makes reference to genetic (within-species), species (between-species) and ecosystem (across groups of species) diversity, which are related and interconnected through ecology and evolution. In this thesis we are principally concerned with species diversity, and genetic and ecosystem diversity to a lesser extent. Subsequent use of the term biodiversity refers to the variability among different species.

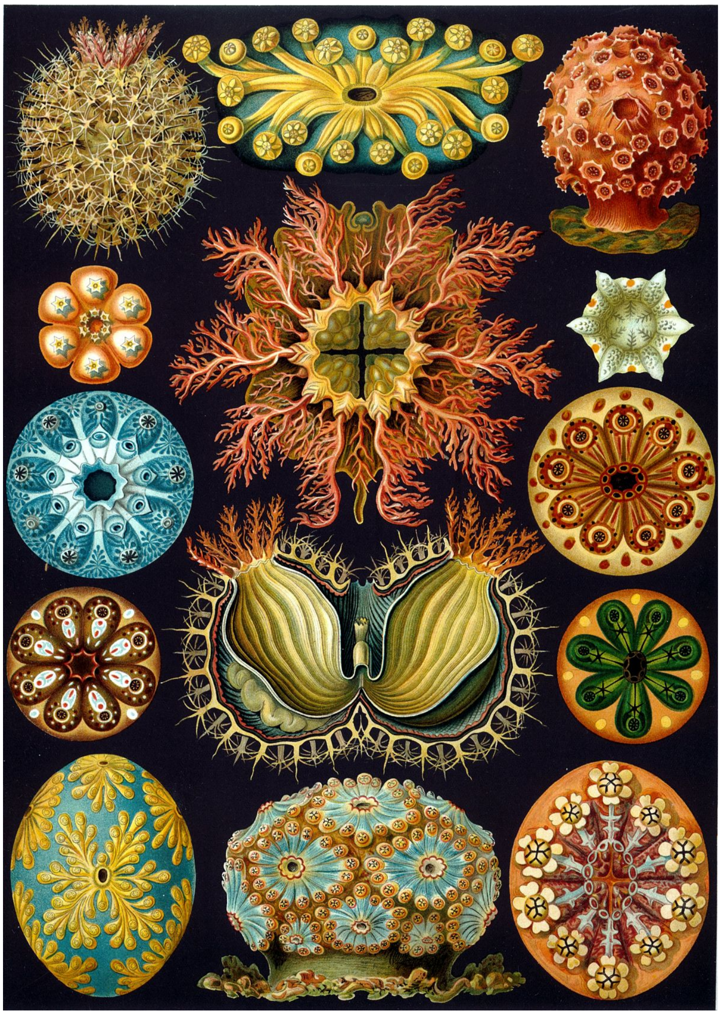


Figure 1 Ascidian biodiversity as shown in Plate #85 of Ernst Haeckel's Art Forms in Nature (Originally Kunstformen der Natur) (Haeckel 1904). *Public Domain*

There have been many recent attempts to estimate the number, or richness of species across the globe (Pimm et al. 2014). The richness of terrestrial (non-bacterial) species has been estimated to be between five and eleven million (Chapman 2009, Costello et al. 2013), with marine estimates somewhat lower at between 0.7-1.0 million (Appeltans et al. 2012). By contrast the richness of global bacterial species, estimated using an extrapolation of species abundance curves, are well over one trillion (Locey and Lennon 2016). This estimate has been criticised (Willis 2016) and subsequent work suggests we should be sceptical concerning reports of trillions of undiscovered prokaryotic species (Schloss et al. 2016). Nonetheless, we can be certain that a vast number of species exist on earth and a large number remain undescribed.

1.1.2 Biogeographic regions & ecosystem services

A second remarkable feature of life on earth is its geographic distribution. The identity of biological forms changes through space, with greater differences between groups of organisms as geographic or environment distance increases (Buckley and Jetz 2008). One of the first dedicated analyses of this pattern was the delineation of 'Zoological Regions' (Wallace 1876), defining six major regions (hereafter realms) still recognised today (see Fig. 2 below). These terrestrial realms are composed of areas of the earth's surface containing broadly similar taxa. Further classifications of realms have involved taxonomic (Proches and Ramdhani 2012) and phylogenetic (Holt et al. 2013) data and divided Wallace's original classification into 11 smaller realms.

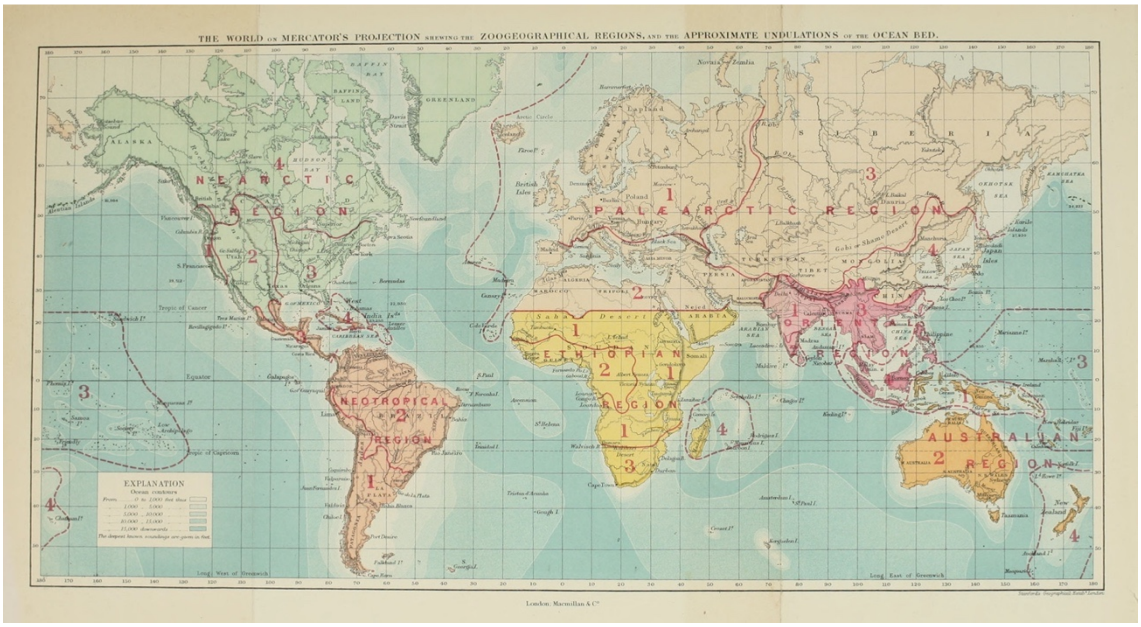

THE WORLD ON MERCATOR'S PROJECTION SHEWING THE ZOOGEOGRAPHICAL REGIONS, AND THE APPROXIMATE UNDULATIONS OF THE OCEAN BED.
NEARCTIC REGION
PALÆARCTIC REGION
ETHIOPIAN REGION
ORIENTAL REGION
NEOTROPICAL REGION
AUSTRALIAN REGION
EXPLANATION
London: Macmillan & Co.

Figure 2 Global map indicating the Zoological Regions from *The Geographic Distribution of Animals* (Wallace 1876). *Public Domain.*

The delineation of biogeographic realms in marine systems has historically received comparatively less attention [but see Ekman (1953)]. However recent work has designated both broad-scale realms (Costello et al. 2017) and ecoregions: a more localised biogeographic region, smaller than a realm (Spalding et al. 2007). These designations find divergence between the different ocean basins, but also uncover variation within basins and proximate to coastlines (see Fig. 3 below).

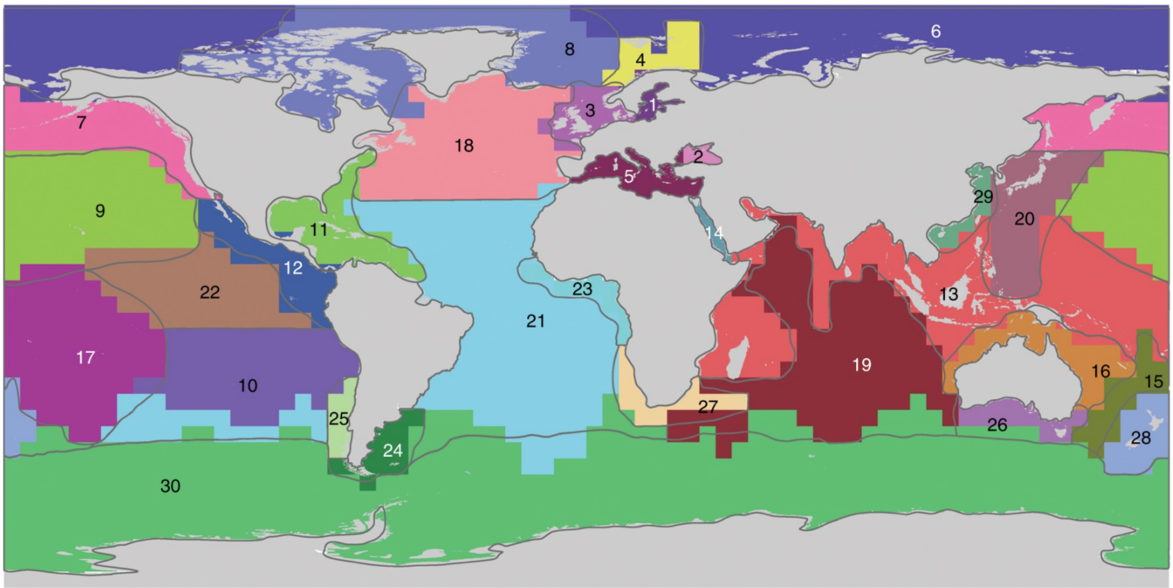


Figure 3 Biogeographic realms identified in Costello et al. (2017) . Adapted from Figure 2a of Costello et al. (2017).

Some key determinants of these global biogeographic boundaries are temperature (both mean and heterogeneity) and tectonic movements or barriers, and both these sets of drivers are important in terrestrial (Ficetola et al. 2017) and marine systems (Costello et al. 2017, Kocsis et al. 2018).

Across geographic space, organisms can occupy many unique functions in the ecosystems they inhabit. Historically, biodiversity research has focussed on numerical changes in species metrics over time (Magurran 2013). However, research has identified that individual species have effects on the physical formation of ecosystems, play fundamental roles in biogeochemical cycling, and occupy key positions in the productivity of biological communities [see Coleman and Williams (2002) for examples]. Together these observations have prompted substantial research into the link between biodiversity and the cumulative function of all species in an ecosystem (ecosystem function) (Gamfeldt et al. 2015, van der Plas 2019). There is now broad consensus across study systems of a non-linear, positive saturating relationship between species richness and ecosystem function (Cardinale et al. 2012, Gamfeldt et al. 2015, van der Plas 2019). Understanding this relationship is important because ecosystem function directly contributes to the ecosystem services required by humanity [marine examples include fisheries, flood prevention and carbon sequestration (Barbier 2017)]. Thus, continued provision of ecosystem function, goods and services relies on understanding, maintaining and protecting biodiversity.

1.1.3 Threats to biodiversity

Biodiversity is in constant flux. Fossil evidence has shown that the richness of animals and plants has dramatically changed through geological time. The trend since the origin of life is an increase in species diversity (Signor 1994), punctuated by a number of major extinction events and subsequent diversification (McElwain and Punyasena 2007, Harnik et al. 2012). A recent analysis of marine microfossils has shown that the bulk of geological time is characterised by relatively stable biological community composition, interspersed by pulses of community change to novel states (Pandolfi et al. 2020). Importantly, shifts to novel community states were correlated with increased local extinction rates, providing some clues from geological timescales into possible outcomes of the current biodiversity crisis.

It is now widely accepted that the Anthropocene is characterised by complex changes in biodiversity. Analyses of species richness trends have shown precipitous global declines over recent decades (Ceballos et al. 2017, Wagner 2020). Conversely, local species richness appears to be increasing through time (Dornelas et al. 2014, Elahi et al. 2015, Vellend et al. 2017) and analyses have not yet found a complete explanation to this apparent discrepancy (McGill et al. 2015, Gonzalez et al. 2016). However, there is greater universal support for elevated extinction and colonisation rates (Dornelas et al. 2019a), compositional turnover (Blowes et al. 2019) and homogenisation of biological communities (Dornelas et al. 2014, Capinha et al. 2015) across Anthropocene ecosystems.

The principal anthropogenic drivers of these observed changes in biodiversity, outlined in the Millennium Ecosystem Assessment (2005), can be categorised into five impact groups that affect marine ecosystems shown in bold below.

It is now widely accepted that **contemporary climate change** has resulted in an increase in a global sea temperatures (Barnett et al. 2001). This change in temperature is linked to a number of other physical changes in the ocean such as ice cover, ocean circulation, thermal stratification and the persistence of upwelling systems (Worm and Lotze 2021). Cumulatively these changes have a direct effect on species, with warm-adapted species replacing cold-adapted ones in biological assemblages [known as topicalization (Vergés et al. 2014)]. This pattern is particularly apparent in marine ecosystems, where species occupy a greater proportion of their possible range extent (Sunday et al. 2012). Indirect effects of climate change, such as trophic changes or the introduction of diseases, have also been shown to have a serious and significant effect on marine biodiversity (Harvell et al. 1999, Worm and Lotze 2021).

Humans also affect marine biodiversity directly through **overexploitation** of marine resources, which has persisted for centuries (Duarte et al. 2020). However, industrialised fishing has had the greatest impact on ecosystems in the last 50 years (Duarte et al. 2020). Coastal fishing has a dramatic effect on food webs (Jackson et al. 2001) and trawling has been shown to reduce benthic diversity and alter community structure (Clark et al. 2016).

Fishing and coastal development are linked, and associated **habitat loss / change** is another group of significant anthropogenic activities. Increased construction of urban coastal infrastructure directly competes with natural systems (Bugnot et al. 2021). In particular, seagrass meadows and mangrove forests have been dramatically affected (Waycott et al. 2009, Polidoro et al. 2010), this is important as these ecosystem engineers create habitat for other marine species (Coleman and Williams 2002).

Alongside the direct effects on ecosystems, coastal urbanisation results in increased **pollution.** In coastal systems, organic enrichment and eutrophication particularly impact on biodiversity (Brauko et al. 2020) and sedimentation has been to shown to have a negative effect on the diversity of phototrophic species, soft-bottom and coral reef communities (Magris and Ban 2019). Finally, the introduction of **invasive species** also has a dramatic effect on biodiversity and is covered in detail below in Section 1.2.

Many of these anthropogenic impacts co-occur in space and time [for example, marine infrastructure construction promotes the invasion of marine species (Airoldi et al. 2015)], and research has identified that these co-occurring threats can be described as a number of spatially separated threat complexes (Bowler et al. 2020). However, the bulk of empirical research into the effects of anthropogenic effects on biodiversity has focussed on single threats, and there is an absence of research into synergistic or interactive effects of multiple drivers of biodiversity loss (Mazor et al. 2018). A meta-analysis of multiple stressors on marine systems consistently found interactive effects across trophic levels (Crain et al. 2008), reinforcing the importance of considering how multiple drivers interact to alter marine biodiversity. Therefore, it is hardly surprising that not a single Aichi biodiversity target from the United Nations 2010-2020 decade of biodiversity was fully achieved (Secretariat_of_the_Convention_on_Biological_Diversity 2020). While there are many reasons for this failure, a comprehensive overview is outside the scope of this thesis. However an important issue in the marine realm is biological monitoring, which poses a consistent problem across different impact types. We rapidly need to improve global coverage and sensitivity of marine biodiversity monitoring tools to detect impacts, monitor management responses and measure the restoration of degraded ecosystems (Scholes et al. 2008, Dornelas et al. 2019b).

1.2 Invasive species

1.2.1 Defining the invader

Scientists have long been interested in anthropogenic movement of animals and plants across the globe. However, terminology and definitions to describe these species have been, and continue to be, subject to intense scientific debate (Richardson et al. 2000, Colautti and MacIsaac 2004, Pyšek et al. 2004, Essl et al. 2018). For example Colautti and MacIsaac (2004) list over thirty different common English terms describing these species, and show that the definitions for each of the terms have distinct scientific meanings in invasion ecology. Here, we define alien, non-native or non-indigenous species as those that have moved from their natural range into a novel region or locality, either directly or indirectly as a result of human action. A subset of the total pool of non-native species develop self-sustaining populations and have a harmful effect on natural ecosystems, human health or local and global economies. This group of non-native species can be classified as invasive species, although more narrow definitions can be applied. For example, the European Union defines an invasive species in regulation 1143/2014 as a species found to threaten or adversely impact on biodiversity and related ecosystem services (EU 2014), making no direct reference to economic impacts. The complexity in defining these organisms is not driven simply by a lack of standardization, there are other complexities that make it hard to provide a universal designation.

One problem lies in the lack of a complete taxonomic description for all organisms, as only a minority of species have been described by science (Appeltans et al. 2012, Costello et al. 2013). Additionally, a large number of described species have been subsequently identified as cryptic; morphologically indistinguishable separate species (Darling and Carlton 2018, Jarić et al. 2019). This leads to uncertainty concerning the native range of species that are only just being described, or in the case of species being reclassified after the discovery of a genetic cryptic linage (Darling and Carlton 2018). Anthropogenic redistribution has resulted in discrepancies between the current and native range of these organisms, and genetic analyses have not been able to resolve the problem because of the reshuffling of linages (Pineda et al. 2011). This makes the reconstruction of the native range for a newly described species difficult, and if a native range cannot be determined, a species cannot be designed alien or native in a given location.

Even for organisms whose taxonomy is resolved, there are complexities. A key issue is the attribution of human influence. Research has documented climate driven range shifts in plants and animals (Walther et al. 2009, Pecl et al. 2017), but some definitions do not consider these species to be non-native (EU 2014) and difficulty arises with examples of organisms responding to human driven climate change with only marginal range shifts (<10 km) per year (Sunday et al. 2012). There is good evidence these species are moving in response to human activity, but they are not routinely designated as non-natives.

Another example can be found in species that have dispersed, adhered to, or associated with rafts such as vegetation or floating natural debris. Marine plastic pollution breaks down more slowly than natural debris, providing a human derived pathway for range expansions across continents (Barnes 2002, Carlton et al. 2017). However, the extent to which these species can be defined as non-native remains unclear, with some authors suggesting these species should be described as neonatives (Essl et al. 2019).

1.2.2 Invasion process

In an effort to understand species invasions, researchers have developed a series of frameworks to categorise species according to the process of a hypothetical invasion. Such an invasion is shown in Fig. 4 below with the stages coloured according to Blackburn et al. (2011).

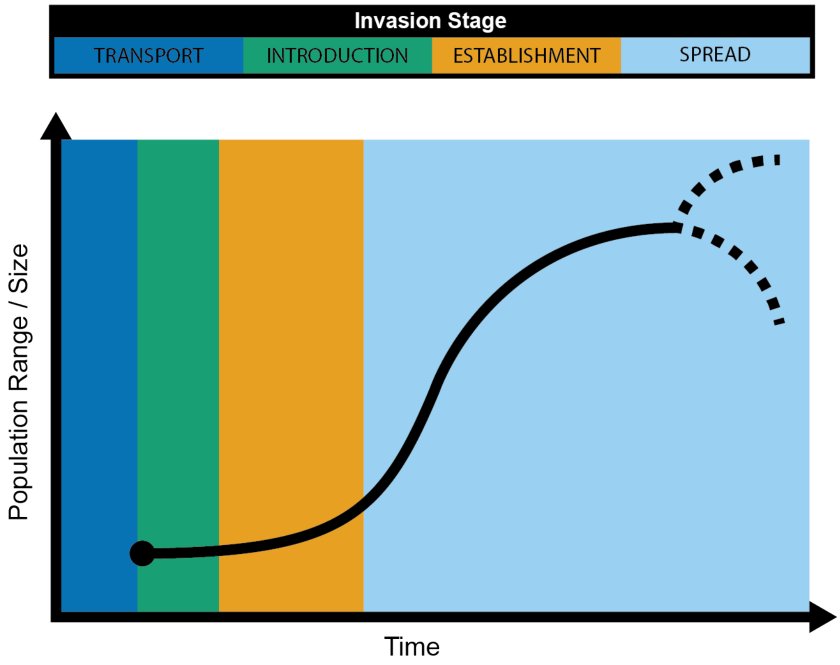


Figure 4 Stylised invasion process showing the advance of an invasive species and the associated invasion stages (Blackburn et al. 2011). The solid line indicates the expansion of range over time with the dashed lines indicating possible outcomes . Adapted from Bourne et al. (2018).

The hypothetical invasion comprises an initial transport stage followed by the introduction into the non-native range. The species may or may not survive the conditions found in the new range. If it does survive and reproduce it enters the establishment stage where it may remain at low abundance for some time. Following the establishment phase some invaders may increase further in abundance and spread to new locations. A number of these frameworks have been proposed (Williamson and Fitter 1996, Richardson et al. 2000, Colautti and MacIsaac 2004, Pyšek and Richardson 2010), with some focussing on barriers between different stages and others prioritising the stages themselves. The unified framework presented in Blackburn et al. (2011) has established a robust set of definitions for each of the stages allowing greater understanding among researchers and between academia and management.

1.2.3 Marine vectors

Central to understanding species invasions is an understanding of the mechanism through which they spread. These pathways, known as vectors, have been the subject of intense study and an overview of terrestrial and freshwater vectors can be found in Hulme et al. (2008). Non-native species in marine ecosystems are predominantly spread as a result of shipping (Williams et al. 2013, Ruiz et al. 2015). Large vessels require ballast to adjust trim and stability, typically vessels pump seawater between onboard ballast tanks and the surrounding environment, which can result in the transfer marine organisms (Bailey 2015). Additionally, sessile organisms can adhere and grow on seawater exposed surfaces of vessels. This process, known as biofouling, has also been shown to be responsible for a large proportion of the total non-native species (Molnar et al. 2008, Williams et al. 2013). While commercial shipping represents the largest biofouling vector by number of vessels and available surface (Williams et al. 2013, Moser et al. 2017), biofouling via recreational vessels have also been shown to contribute to species invasions (Davidson et al. 2010, Ashton et al. 2014). After shipping, the release of species associated with aquaculture is the next largest vector by number of non-native species (Molnar et al. 2008, Williams et al. 2013). In these cases adult organisms can be released unintentionally [for example the escape of Atlantic Salmon as a result of a storm (Jensen et al. 2010)] or gametes, planktonic or juvenile organisms may be unintentionally released into the natural environment (Diederich et al.

2005). Additionally, the movement of aquaculture species during their lifecycle, such as transport from the hatchery to the ocean, represents a vector as fouling species are often found adhered to shellfish (Grosholz et al. 2015). The opening of canals is a major vector for the spread of non-native marine organisms and in some regions contributes a greater proportion of non-native species than aquaculture as shown in Fig. 5 below (Katsanevakis et al. 2013). Finally, there are several vectors that contribute a small number of non-native species such as the live seafood and bait trades (Weigle et al. 2005), the release of ornamental organisms (Padilla and Williams 2004) and the release of live animals for religious reasons (Magellan 2019).

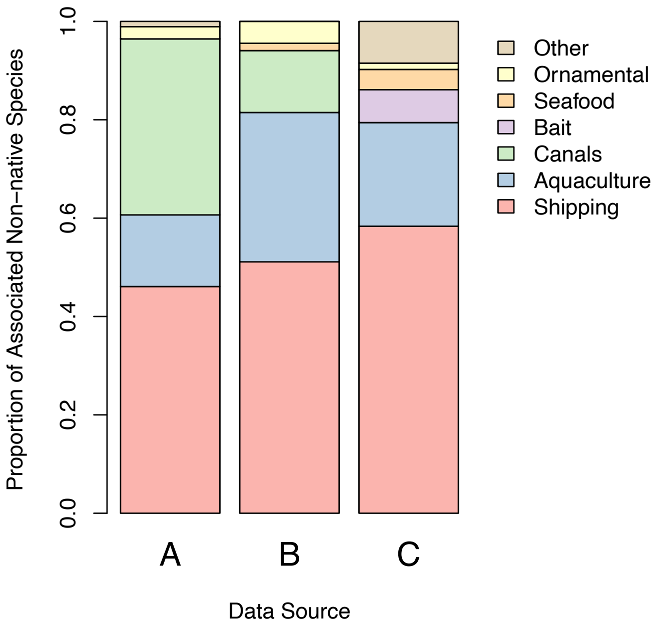


Figure 5 Stacked bar chart showing vector associated with non-native species introductions. Data sourced from **A** Katsanevakis et al. (2013) **B** Molnar et al. (2008) **C** Williams et al. (2013).

1.2.4 Marine impacts

Studies estimate that around 10% of known non-native species have a negative effect on the ecosystem into which they are introduced (Ricciardi and Kipp 2008, Anton et al. 2019). In many cases there is a lag between the initial introduction and the detection of increases in population density or range (Azzurro et al. 2016). Despite a large proportion of non-native species that have no detectable impact, there is substantial evidence that those that do have significant and wide-ranging effects on both biological and human systems.

The economic costs of invasive species are estimated to be billions of USD per year per country (Marbuah et al. 2014, Diagne et al. 2021). These costs can be direct, such as those associated with removing fouling species which affect the efficiency of vessels during transit (Fernandes et al. 2016). Alternatively, these costs can reduce the value of existing resources such as fisheries, for example the European Green Crab (*Carcinus spp.*) has impacted clam fisheries in both Tasmania (Walton et al. 2002) and Maine (McClenachan et al. 2015), and there is evidence it has a negative effect on other fisheries across their invaded range (Leignel et al. 2014). These estimated costs rarely cover the negative impacts on ecosystem services (Vilà and Hulme 2017) as the effect of invasion can be hard to assess in financial terms, even when the monetary value of marine ecosystem services is known (Costanza et al. 1997).

A number of meta-analyses have evaluated the effect of invasive species on native species, communities and ecosystems (Thomsen et al. 2014, Gallardo et al. 2016, Anton et al. 2019). All studies found large variation in the size of the impact of invasive species across trophic levels and impact type, demonstrating that the outcome of an introduction is context specific. Furthermore, invader trophic level has been shown to be an important factor, with predators having a strong negative effect on biodiversity and ecosystem properties, with mixed effects at lower trophic levels (Gallardo et al. 2016, Anton et al. 2019). While a global examination of the effects of all invasive species necessarily ignores the context of the individual invasions, in general species invasions were associated with a modest but significant decrease in ecological variables (such as water turbidity or sedimentation rate) (Gallardo et al. 2016, Anton et al. 2019).

Synthesising multiple species invasions to produce generalities across trophic levels is difficult, although there is evidence that many invasive species have negative effects on ecological systems (Simberloff et al. 2013). Furthermore, a global analysis recently identified that pre-existing anthropogenic impact may have a positive interactive effect on ecological changes due to species introductions (Geraldi et al. 2020), further complicating the measurement of impact.

1.2.5 Current issues & opportunities

There are a number of pertinent issues concerning non-native species in marine systems. First, transoceanic shipping is projected rise across the century, increasing the global invasion rate 3-20 fold (Sardain et al. 2019). Additionally, there are a large number of planned shipping infrastructure improvements, such as the expansion of the Suez (Galil et al. 2015) and Panama canals (Muirhead et al. 2015), which will increase invasion potential. Moreover, large international development programs, such as the People's Republic of China's Belt and Road Initiative plan to create transport super-corridors across the globe, will break down biogeographic boundaries promoting species invasions (Liu et al. 2019). Contemporary climate change has also had an effect on the global expansion of international shipping with new routes in the Arctic region now possible due to the loss of sea ice (Miller and Ruiz 2014). There have been a growing number of non-native species introductions into the Arctic in recent decades, with most of the introductions attributed to shipping (Chan et al. 2019). These findings indicate that the relatively pristine Arctic region is likely to be impacted by an increasing number of invasive species in the near future, and sets a precedent for other remote polar regions such as Antarctica (McCarthy et al. 2019). Another concern is the impact of microbial species invasions. A great deal of research has identified that pathogenic microorganisms can cause substantial shifts in ecosystems across marine trophic levels (Bourne et al. 2009, Hewson et al. 2014). Comparatively little is known about global patterns of marine microorganismal biogeography, with wide-scale studies limited to sample sizes capable of picking up only granular biogeographic patterns (Sunagawa et al. 2015, Richter et al. 2019). We therefore have a limited understanding of what constitutes a native or non-native range for these species and have even less data concerning the frequency and severity of micro-invasions.

Meeting these challenges in the coming decades will require new approaches and tools that enable worldwide detection of marine invaders at a fraction of current costs. The use of citizen science in biodiversity research has expanded our ability to monitor ecosystems

(Theobald et al. 2015) and examples of marine invasive species being monitored using citizen science programs are becoming more common (Scyphers et al. 2015, Giovos et al. 2019). While data quality may in some cases be lower than that collected by professional scientists, the reduction in costs, increase in spatial extent and number of records that can be generated make it a useful tool for marine invasions. Most pertinent to this work is the use of genetic methods for the early detection of marine non-native species across the tree of life.

1.3 Environmental DNA

1.3.1 What is eDNA?

Environmental DNA (or eDNA) is genetic material lost by an organism into the environment in which it lives, and subsequently isolated from environmental samples such as water or sediment. It is typically a mixture containing many different types of biologically derived particles all of which contain polynucleotide molecules (Fig. 6).

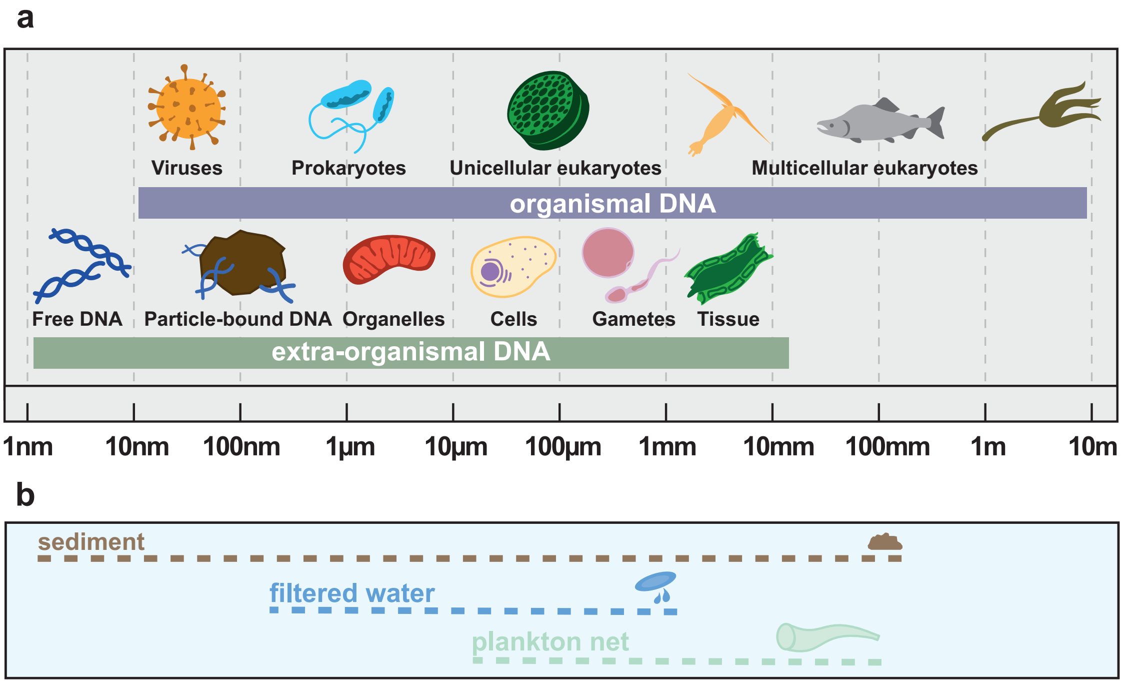

a
Viruses
Prokaryotes
Unicellular eukaryotes
Multicellular eukaryotes
organismal DNA
Free DNA
Particle-bound DNA
Organelles
Cells
Gametes
Tissue
extra-organismal DNA
1nm
10nm
100nm
1µm
10µm
100µm
1mm
10mm
100mm
1m
10m
b
sediment
filtered water
plankton net

Figure 6 Conceptual diagram of the sources of eDNA showing **a** the overlapping size distribution of extra-organismal and organismal DNA, and examples of various eDNA components. **b** the size distribution captured by various tools typically used in eDNA studies. Figure from Rodriguez-Ezpeleta et al. (2021).

While it has been known for some time that DNA isolated from the environment can reflect the biodiversity of the region from which the sample originated, the first use of eDNA to study contemporary metazoans using water samples was Ficetola et al. (2008) [but see Martellini et al. (2005)]. This study prompted a rapid expansions of eDNA research (see Fig. 7 below), the publication of a textbook (Taberlet et al. 2018), an eDNA specific journal (Bernatchez 2019) and the formation of an academic society (Minamoto et al. 2020). Many commercial providers and natural resources agencies across the globe have also begun to integrate genetic data into their biodiversity assessments.

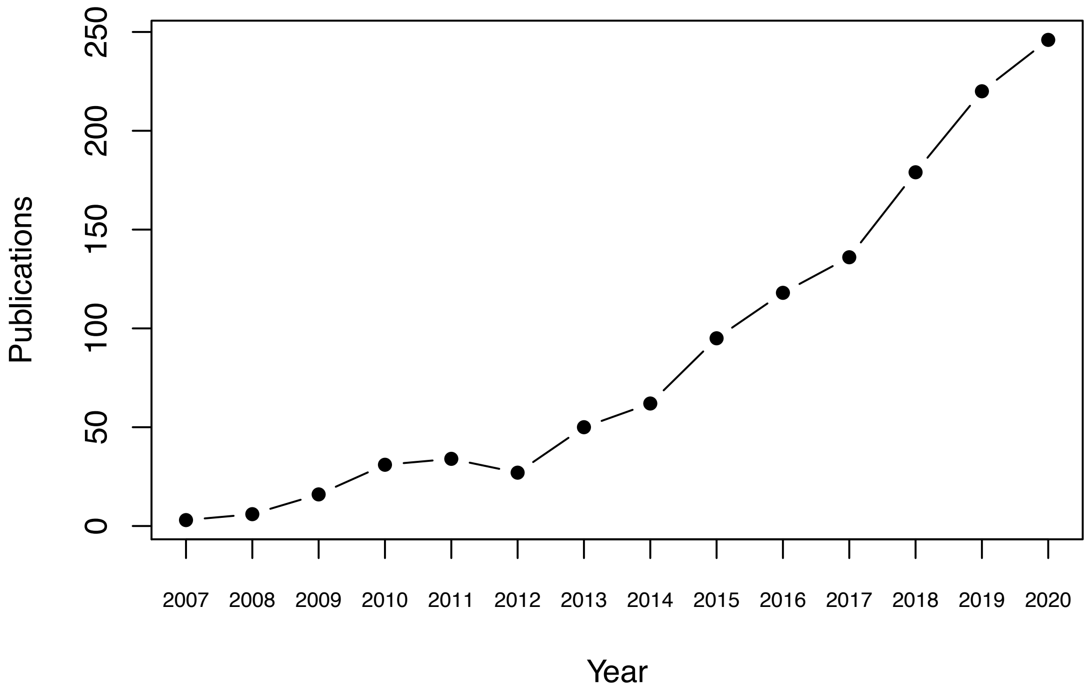


Figure 7 Number of publications with the term 'eDNA' found in the title or abstract per year in the NCBI Entrez database as of 01/02/2021. Data not collected for 2021.

1.3.2 The ecology of eDNA

As eDNA biodiversity surveys have become more common, the behaviour of individual eDNA particles through time and space has become an important subject of study. If a researcher detects an DNA molecule corresponding to an organism of interest what does that say about the incidence of the organism in the local area, and how should she interpret a positive result? The study, or ecology, of eDNA (*sensu* Barnes and Turner) aims to describe the processes that results in eDNA being detectable in a given environment; the **origin** of eDNA, the persistence or **state** of the particles, the ways in which eDNA is **transported** actively or passively through differing environments and the ultimate **fate** of eDNA. Through understanding each of these processes we learn to interpret the presence, or absence, of eDNA in our environment of interest, and how this links to the biodiversity of the ecosystem. The following sections overview research into these four 'Ecology of eDNA' processes.

1.3.2.1 Origin

There is evidence that the majority of eDNA comes from life processes such as egestion and excretion, either directly into the environment or via eDNA transport (Wotton and Malmqvist 2001, Caldwell et al. 2011, Williams et al. 2015). Beyond these processes there is likely vast interspecific variation in the origin of eDNA from other sources. Some organisms (for example, teleost fish or molluscs) produce mucus that performs a variety of functions from immunity to physiological homeostasis (Shephard 1994, Davies and Hawkins 1998). Studies have shown that this mucus contains quantities of DNA sufficient for genetic analysis (Henley et al. 2006, Le Vin et al. 2011, Taslima et al. 2016), suggesting mucus as a source of eDNA for these organisms. By contrast, many marine organisms rely on a hard exoskeleton which might precent continuous loss of DNA from the outer surface of these organisms. DNA has, however, been isolated from exoskeleton remains after moulting (Bista et al. 2017) so this might be a source of eDNA in these cases. A number of studies have isolated eDNA attributed to spawning events (Bayer et al. 2019, Ratcliffe et al. 2021), though it remains unclear if these observations represent cellular debris, gametes or entire larvae captured during filtration. Finally, there is mixed evidence for the contribution of carcasses to eDNA in water bodies, with studies finding that fish carcasses produce detectable levels of eDNA (Merkes et al. 2014, Kamoroff and Goldberg 2018) while no detectable eDNA was found to be produced by crayfish carcasses (Curtis and Larson 2020).

The quantity of eDNA produced in metazoans is affected by a range of abiotic and biotic factors, reviewed in Stewart (2019). The rate of eDNA shedding has been shown to increase with water temperature in most fish species (Lacoursière-Roussel et al. 2016, Robson et al. 2016, Jo et al. 2019b), but evidence suggests that this effect is not universal (Klymus et al. 2015). Mesocosm experiments show that greater organismal biomass results in a larger concentrations of eDNA (Takahara et al. 2012, Maruyama et al. 2014, Jo et al. 2019b) and that for a given biomass juvenile individuals produce more eDNA per biomass than adults (Maruyama et al. 2014). Additionally, eDNA shedding may be higher during behaviours such as feeding (Klymus et al. 2015, Ghosal et al. 2018) or fighting (Dunn et al. 2017) but a full manipulative experiment separating out the effects of each behaviour, intraspecific competition and the movement associated with the behaviour has not yet been attempted. Thus, growing evidence indicates that the production of eDNA varies both within and between species.

Depending on the isolated environmental substrate, eDNA may additionally contain whole live organisms such as planktonic juveniles or microscopic meiofauna. Researchers typically make no distinction between DNA from these sources, likely because it would be impossible to do so. Yet, some researchers consider bulk DNA sample collection a different approach in which live organisms are collected or trapped and processed for DNA extraction together [reviewed in van der Loos and Nijland (2020)]. These bulk DNA samples are still eDNA in that they are bulk DNA extracted from the environment, but the as the whole organism has been collected the statistical assumptions are different (Rodriguez-Ezpeleta et al. 2020). Examples of bulk organismal DNA collection methods include plankton nets, terrestrial pitfall traps and settlement plates.

Once shed, eDNA has been isolated from different substrates including water (Deiner et al. 2016), sediment (Koziol et al. 2019), sand (Naviaux et al. 2005), honey (Utzeri et al. 2018), faeces (Carroll et al. 2018), flowers (Thomsen and Sigsgaard 2019) and blood (Schnell et al. 2018). However the majority of work in metazoans has focussed on either water or sediment (Jarman et al. 2018).

1.3.2.2 State

The state of eDNA in different substrates is important, as different types of eDNA particle degrade at variable rates. Understanding eDNA state is also useful to target particles of interest such as identifying freshly produced eDNA to detect a rare species or ancient eDNA for palaeoecological studies. Experiments passing aqueous eDNA through columns packed with different substrates have shown that eDNA is polydisperse (contains a number of different sized particles) and heterogenous (has a variable concentration even when mixed) (Shogren et al. 2016). Further studies performing sequential filtration on aqueous eDNA have shown that eDNA in fish is predominantly between 1-10 µm in size (Turner et al. 2014, Wilcox et al. 2015, Jo et al. 2019a), while some evidence suggests that eDNA in invertebrates may be much smaller (<1 µm) (Moushomi et al. 2019).

In sediment, studies have shown that individual DNA molecules can bind directly to clay (Pietramellara et al. 2007) or sand particles (Lorenz and Wackernagel 1987). This can protect the DNA from degradation until it is released from particles through DNA extraction (Naviaux et al. 2005, Pedersen et al. 2016).

1.3.2.3 Transport

Once eDNA is present in the environment transport mechanisms can be categorised into two broad groups: horizontal and vertical. Horizontal eDNA transport refers to mechanisms that move eDNA across the land or seascape. Evidence shows that eDNA in lotic systems can be transported many kilometres from the source (Deiner and Altermatt 2014), but evidence in non-flowing systems suggests that eDNA may remain closer to its source (Li et al. 2019, Murakami et al. 2019). Vertical transport refers to the processes that result in eDNA becoming fixed in freshwater sediments or marine benthos (Wotton and Malmqvist 2001, Turner et al. 2015). For example, eDNA from vertebrates species settling in pond sediment (Buxton et al. 2018). Vertical transport can also resuspend eDNA previously stored in sediments (Shogren et al. 2017), potentially resulting in detections after an organism is no longer present in the environment. These processes demonstrate the need for an understanding of the hydrology of the ecosystem of study in eDNA surveys.

It is possible to take advantage of known eDNA transport mechanisms in biodiversity surveys, for example in the concentration of eDNA from a number of upstream environments (Deiner et al. 2016), or the accumulation of eDNA in lake sediments from terrestrial organisms living in the catchment (Ficetola et al. 2018). While a number of studies have taken advantage of this 'conveyor belt' effect, we lack a comprehensive understanding of the interactions between environmental substrates and eDNA and have limited knowledge about the drivers of eDNA transport.

Trials using experimental river systems have shown little effect of riverbed substrate type on eDNA transport (Jerde et al. 2016), with most variation in eDNA detection remaining unexplained. Further experiments have shown that eDNA substrate type and stream steepness can affect the resuspension of eDNA from the substrate over time (Shogren et al. 2017, Fremier et al. 2019). These studies aside, we still have limited knowledge of the rate of eDNA transfer between different substrates (for example, water to sediment), though work has begun to explore the differences in eDNA concentration and decay between different substrates (Turner et al. 2015, Sakata et al. 2020). Yet despite our limited knowledge of vertical eDNA transport, researchers have found that eDNA from substrates in which organisms live (such as seawater) predominantly reflects contemporary species presence (Bohmann et al. 2014, Deiner et al. 2017).

Substrates that act as sinks for eDNA (such as sediment or ice) can contain information about both contemporary and past biodiversity (Pedersen et al. 2015). Due to eDNA being preserved in stratified layers in some sedimentary environments the interpretation of eDNA from sediment cores requires an established chronology and evidence that the eDNA has not moved between the layers (Haile et al. 2007). This ensures vertical transport has been consistent and ancient eDNA isolated from each strata reflects the biodiversity associated with the period in which the layer formed.

1.3.2.4 Fate

Once cellular material is lost to the environment eukaryotic cell death proceeds either through the initiation of apoptosis or by necrosis, both of which lead to the breakdown of organelles and release of genetic material (Hotchkiss et al. 2009). DNA no longer contained within a cell can be broken down by DNases associated with microorganisms in the environment and incorporated into the carbon, nitrogen or phosphorus cycles (Paul et al. 1987, Pietramellara et al. 2009). DNA also spontaneously degrades in the absence of enzymes (Lindahl 1993); both heat and UV can cause the structure of DNA to change through a number of chemical pathways (Sancar and Sancar 1988). Biological (enzyme driven) degradation of DNA has been shown to occur in a matter of hours (Salter 2018), while spontaneous DNA degradation has a half-life between tens and thousands of years, depending on the chemical pathway and temperature (Schroeder and Wolfenden 2007). Therefore, the most relevant pathway for degradation of eDNA, and of a measurable eDNA signal, is biological. Experiments have consistently shown that UV light has little effect on eDNA detection (Andruszkiewicz et al. 2017, Mächler et al. 2018). However, no study has systematically determined the relative effects of biological or chemical degradation pathways, and research into the effects of bacteria on eDNA decay have shown mixed results (Tsuji et al. 2017, Salter 2018, Zulkefli et al. 2019). Further work is required to definitively determine the comparative importance of different mechanisms of DNA degradation.

A growing number of studies have investigated the effects of various physical, chemical and biotic factors on the decay of detectable eDNA in aquatic systems (Andruszkiewicz et al. 2017, Jo et al. 2019b). They demonstrate that some variables, such as pH or temperature, appear to have a strong effect on eDNA decay (Eichmiller et al. 2016, Tsuji et al. 2017, Seymour et al. 2018, Jo et al. 2019b). As many of these factors vary seasonally, we might expect seasonality to effect eDNA decay, but data in temperate ecosystems has yet to demonstrate a consistent pattern (Collins et al. 2018). Understanding the relationship between these factors and decay is important to ensure appropriate metadata is collected and to avoid false-negative errors. Most marine studies show that eDNA in seawater becomes undetectable after two to five days (Sassoubre et al. 2016, Andruszkiewicz et al. 2017, Minamoto et al. 2017, Collins et al. 2018, Cowart et al. 2018, Jo et al. 2019b, Holman et al. 2021a). However, in some cases positive detections become inconsistent after only 48 hours (Forsström and Vasemägi 2016, Sassoubre et al. 2016). This research reasserts the importance of understanding the eDNA decay rate for each species in a given spatiotemporal environment to provide inference into the relationship between contemporaneous eDNA and species presence.

1.3.2.5 Limitations of the ecology of eDNA

Our understanding of eDNA has advanced since the ‘Ecology of eDNA’ framework was introduced in Barnes and Turner (2016). There are, however, some problems that continue to arise in eDNA research. Firstly, a large proportion of the eDNA work is conducted in fish (Jarman et al. 2018). This may be due to their commercial importance, their availability for experimental studies, their ecological importance in ecosystems or more simply because there are a large number of researchers working on fish species across ecosystems. Regardless of the reason, this is problematic because studies have identified that eDNA dynamics (such as shedding rate or decay) vary across taxa (Collins et al. 2018, Stewart 2019), and some initial evidence suggests that even the particle size distribution of eDNA may vary across species (Moushomi et al. 2019). It is therefore essential, that basic research is conducted into the properties of eDNA for a range of taxa and that findings from different organisms are not generalised across species.

Secondly, it remains unclear to what extent results in tank trials apply in natural systems. A large number of studies are conducted in laboratory aquariums (Sassoubre et al. 2016, Collins et al. 2018, Jo et al. 2019a, Jo et al. 2019b, Moushomi et al. 2019), with a smaller number in outdoor experimental pools or streams more closely resembling field conditions (Seymour et al. 2018, Li et al. 2019). Evidence reviewed above shows that conditions and animal behaviour can affect patterns of eDNA production and decay (Klymus et al. 2015, Dunn et al. 2017, Stewart 2019). Furthermore, we have yet to describe any consistent functional relationship between the ecology of eDNA and bacterial abundance or diversity. Some studies have linked bacterial processes to various components of eDNA ecology (Tsuji et al. 2017, Salter 2018, Zulkefli et al. 2019), but bacterial diversity is rarely listed among required metadata for eDNA studies (Goldberg et al. 2016). If bacteria have a strong effect on the ecology of eDNA, for example through absorbing or digesting free DNA, then the uncontrolled microbial differences between tank trials and unobserved microbial variance in field studies might be an essential unrecorded parameter driving experimental outcomes.

1.3.3 eDNA detection methods

Once extracted from a sample environmental DNA can be stored and assayed as any DNA sample. Two broad groups of analytical approaches have become common in eDNA experiments. One group (single-species approaches) targets a single DNA region in the extracted DNA pool. Providing information about the incidence, and in some cases quantity, of DNA molecules originating from a particular single species. A second group (multi-species approaches) target the DNA molecules from multiple species to determine incidence, and less commonly quantity, in parallel either by targeting a homologous gene fragment or by examining all DNA molecules found in the DNA mixture. An overview of the technologies used in each of the methods is presented below.

1.3.3.1 Single-species targeted detection

Analytical approaches detecting a single species from eDNA samples, sometimes called targeted eDNA analysis, have taken advantage of pre-existing resources, such as DNA barcodes and previously validated primers, to develop assays that enable a researcher to confirm the incidence of a DNA template in an eDNA sample. The bulk of eDNA research to date has been conducted using quantitative polymerase chain reaction (qPCR) (Tsuji et al. 2019, Thalinger et al. 2021) wherein the product of a PCR is quantified over many cycles to determine the quantity of the DNA template in original sample. A large number of studies have confirmed that, with appropriate quality control and validation (Goldberg et al. 2016), the quantified concentration of eDNA from qPCR reflects that in both the DNA extract and in some cases organismal biomass (Plough et al. 2018, Yates et al. 2019). While most studies rely on qPCR instruments, others have shown that visualising PCR products using gel electrophoresis can also provide sensitive detection (Clusa et al. 2017, Blackman et al. 2020). Recent studies have begun to implement digital droplet PCR (ddPCR) in eDNA surveys (Baker et al. 2018, Uthicke et al. 2018). In this method PCR reactions are miniaturised and assayed individually to produce many thousands, or hundreds of thousands, of replicate PCR reactions that provide a binary detection result. The original DNA template concentration correlates with the proportion of binary PCR positives to give an accurate and sensitive quantification of DNA template. Studies have shown this technique to provide more sensitive detection, while being less sensitive to PCR inhibition, albeit at an increased cost per assay, compared to qPCR (Doi et al. 2015, Mauvisseau et al. 2019). The initial design of a qPCR assay is relatively simple, but validation and optimisation can be expensive and difficult. A recent effort, building on previous work (Goldberg et al. 2016), aims to classify eDNA assays based on their level of validation and provides a roadmap for those aiming to produce a reliable eDNA assay for PCR, qPCR or ddPCR platforms (Thalinger et al. 2021). Additionally, some eDNA studies have developed methods optimised for portability and have demonstrated proof of concepts for on-site eDNA detection of species using PCR-based lateral flow assays (Doyle and Uthicke 2020) and loop-mediated isothermal amplification (Williams et al. 2017). These platforms require either a thermo-cycler or rigorous optimisation. An exciting new development that overcomes these challenges is the use of the CAS12a enzyme for nucleic acid detection (Chen et al. 2018, Gootenberg et al. 2018). CAS12a is an enzyme associated with the CRIPSR/Cas system (clustered regularly interspaced short palindromic repeats/CRISPR associated nucleases) predominantly used in gene editing. A limited number of studies have demonstrated this technology on eDNA samples (Williams et al. 2019) showing excellent sensitivity and potential for simple and portable eDNA detection of single species in the near future.

1.3.3.2 Multi-species passive detection

The second group of approaches (passive eDNA analysis) rely on the use of high-throughput sequencing technologies. Broadly, these technologies enable researchers to sequence many millions of DNA molecules in parallel [see Goodwin et al. (2016) for a comprehensive technology review] and have been instrumental in many recent discoveries across biodiversity and invasion ecology (Bourne et al. 2018, Taberlet et al. 2018). Metabarcoding is the most common type of multi-species method employed in eDNA surveys (Deiner et al. 2017, Jarman et al. 2018). This method uses primers that target a homologous short DNA fragment (typically 100-400bp in length) in a PCR that anneals oligonucleotides onto the biologically derived DNA for sequencing. These primers are sometimes called ‘universal’ or ‘degenerate’ indicating that they have binding affinities for a range of different DNA templates. This is usually achieved with primers incorporating one or more degenerate bases (for example, W or M in the International Union of Pure and Applied Chemistry degenerate base symbols table) which is synthesised as a mixture of different oligonucleotide molecules. Alternatively, an deoxyinosine base can be used which can bind to all four bases of the genetic code (Rossolini et al. 1994). Successful metabarcoding relies on sufficient reference data to guide primer design to ensure maximum amplification across the target group, a region being amplified that contains enough genetic variation to delineate different species and minimal cross- amplification with non-target groups (for example, bacteria or protists for a set of primers targeting fish). No commonly used primers satisfy all these conditions perfectly (Leray et al. 2013), but there is minimal cross amplification when the target taxonomic group is relatively restricted (Miya et al. 2020). It is therefore necessary to have an understanding in advance of the primer binding affinity to a given DNA template for correct interpretation of positive and negative detections (Kelly et al. 2019).

A less common technique for eDNA surveys is shotgun-metagenomics. Here a DNA template isolated from an environmental sample is used to build a DNA sequencing library without any specific amplification of any genomic region. While this technique is common in the study of microorganisms (Quince et al. 2017) most metagenomic eDNA studies in larger organisms have been conducted using sedimentary ancient environmental DNA (Pedersen et al. 2013, Pedersen et al. 2015, Pedersen et al. 2016) and very few studies have used it for eDNA biomonitoring (Stat et al. 2017). Stat et al. (2017) showed that shotgun-metagenomics identified a much smaller fraction of the metazoan biodiversity compared to multiple metabarcoding primer sets. It should be noted however that current metazoan metagenomics workflows are less well developed compared to metabarcoding and appropriate quality control metrics, such as number of reads per sample, are not well understood.

One metagenomic technique that has great potential is hybridisation-based approaches. It is possible to use oligonucleotides designed using a reference library containing target species to enrich a DNA sample by hybridising environmental DNA to the synthetic references (Wilcox et al. 2018). Once enriched, the DNA can then be shotgun sequenced. Early studies have shown this approach detects a greater number of species compared to metabarcoding (Giebner et al. 2020) or shotgun-metagenomics alone (Murchie et al. 2020). However, this technique is comparatively expensive, requires complete DNA reference sequences for detection and requires large DNA input under some protocols (Wilcox et al. 2018, Giebner et al. 2020).

1.3.4 eDNA in the oceans

Despite their scale and global importance, a relatively minor proportion of eDNA biodiversity studies have been conducted in marine systems [see Fig. 2 from (Beng and Corlett 2020)]. We are only beginning to appreciate how eDNA surveys can contribute to marine science.

Observational ecological datasets covering a large number of oceanic basins and ecosystems have begun to describe the structure and function of biological communities (Sunagawa et al. 2015, Richter et al. 2019) revealing complex patterns driven by abiotic variables and species interactions (Djurhuus et al. 2020) across the tree of life (Holman et al. 2021b). Additionally, eDNA has been used in field manipulation studies (Nascimento et al. 2019) allowing researchers to describe treatments with greater accuracy and resolution compared to existing biodiversity survey tools.

Beyond fundamental ocean science, eDNA also has a use in conservation. Studies have shown that eDNA can accurately detect rare, mobile species of conservation concern, such as sharks (Boussarie et al. 2018) or marine mammals (Djurhuus et al. 2020). It is also possible to establish seasonal residency of marine organisms with a view to advising conservation efforts (Stoeckle et al. 2017, Postaire et al. 2020). Beyond single species assessments, eDNA is expected to contribute to all stages of marine spatial planning [See Bani et al. (2020) for a review] from initial surveys, in the determination of ecosystem status (Stat et al. 2017) and to evaluating the health of marine protected areas (Gold et al. 2021) and fisheries post-designation (Salter et al. 2019).

A number of studies have explored the utility of eDNA in marine fisheries science with experiments comparing eDNA to acoustic (Yamamoto et al. 2016), visual (Sigsgaard et al. 2017) and trawling data (Knudsen et al. 2019, Salter et al. 2019). They indicate that eDNA is a suitable tool for fisheries management, but standardisation is required for specific fisheries to ensure comparability with existing methods (Jerde 2019, Salter et al. 2019). Aquaculture has also begun to incorporate methods based on the collection and analysis of eDNA. Studies have shown that sediment eDNA collected from around open aquaculture pens in marine systems can be analysed to assess the impact of production on the environment (Pawlowski et al. 2014, He et al. 2019, Frühe et al. 2020). eDNA can also be used to detect bacteria or parasites associated with fish disease in aquaculture facilities (Peters et al. 2018, Fossøy et al. 2020), enabling treatment or early harvest to prevent financial losses. Finally, recent work has identified that non-invasive sampling of valuable adult breeding organisms can be achieved with eDNA (Espinoza et al. 2017, Holman et al. 2019b) avoiding mortality and handling stress associated with DNA biopsies.

The use of eDNA in marine invasion ecology has increased in recent years (Duarte et al. 2021) with studies demonstrating the advantages of molecular approaches across the stages of invasion. Experiments have shown that non-native species can be detected during transport using eDNA, for example in ballast water (Gerhard and Gunsch 2019, Rey et al. 2019). Work has also shown that species recently introduced into harbours and marinas can be detected using eDNA metabarcoding (Deiner et al. 2018, Grey et al. 2018, Holman et al. 2019a, Westfall et al. 2020) at lower cost than existing methods (Borrell et al. 2017). Studies have demonstrated that a management response is more likely to be successful and can cost less during the early stages of an invasion (Leung et al. 2002, Beric and MacIsaac 2015) showing the importance of eDNA detection in the transport and introduction stages of invasion. However, some studies have produced false negative detections of non-native species using qPCR of water eDNA samples (Forsström and Vasemägi 2016, Wood et al. 2018) emphasising the importance of understanding eDNA transport and assay sensitivity for novel species. As invasions proceed into the 'establishment' and 'spread' phases, eDNA can be used to detect further range expansions (Dufresnes et al. 2019, Carim et al. 2020), ensure eradication efforts have been successful and in future monitoring to check for reinvasion (Dunker et al. 2016, Miralles et al. 2016, Davison et al. 2017). Finally, recent work has identified that ancient sedimentary eDNA can be used to reconstruct pre-invasion community biodiversity information (Ficetola et al. 2018, Nelson-Chorney et al. 2019). In summary eDNA-based methods revolutionise our ability to understand our oceans, reveal changes in marine community composition in response to anthropogenic and natural drivers, and enable a methodological step-change in several avenues of applied sciences such as fisheries and invasion ecology.

1.4 Thesis aims & overview

This thesis asks: *how can eDNA contribute to our knowledge of marine biodiversity and biological invasions?* Each chapter is presented in the form of a publication and, where appropriate, the peer-reviewed publisher typeset article is presented along with supplementary information.

Chapter Two asks if eDNA metabarcoding can be used to detect novel and previously reported marine species invasions, and how methodological choices affect the results. Replicate water and sediment samples from four recreational marinas across the United Kingdom were collected. These samples were subjected to eDNA metabarcoding of a cytochrome c oxidase subunit I (COI) and nuclear small subunit ribosomal DNA (18S) gene fragments. The resultant dataset was compared to biodiversity survey data collected by taxonomic experts, and the overlap of detected species was detailed.

Chapter Three asks if it is possible to use the developed tools and eDNA methods detailed in the previous chapter to profile biodiversity in Anthropocene oceans in a standardised way. To do this water samples from 18 varied sites were collected to represent the full range of ecoregions across the South African coastline, cumulatively stretching over 2000km. Environmental DNA was extracted from these samples and subjected to multi-marker metabarcoding targeting organisms spanning many orders of magnitude in organismal size across the tree of life, from bacteria to metazoans. First, broad-scale biogeographic patterns for each taxonomic group were analysed. Environmental variables were then used to explain the observed community dissimilarity. Finally, evidence for homogenisation of communities across taxonomic groups was evaluated .

In **Chapter Four** examines the ability for eDNA to detect changes in the invaded range of four sessile non-native marine organisms over a period of eight years, evaluates changes in genetic diversity across this period, and explores possibilities for eDNA to be used for population genetics. Biodiversity survey data from 2009 was compared to newly collected rapid assessment and eDNA metabarcoding biodiversity survey data on the abundance and distribution of four non-native ascidian species (Class *Ascidiacea*) across the coastline of South Africa. Tissue samples were collected for the same four species and a region of the cytochrome c oxidase subunit I gene (COI) was sequenced for each sample.

Chapter Five is a study conducted as a part of a CASE studentship (Collaborative Awards in Science and Engineering) with commercial partners Xelect Ltd. In this chapter a non-invasive genotyping technique using eDNA is developed with an aquaculture application. The method demonstrates the collection of a DNA sample from small or valuable organisms for which tissue sampling is not possible or preferable, as in many marine biodiversity or invasion biology studies. The method was developed for a bivalve mollusk and relied on the release of eDNA by the organisms into a small volume (500ml) of salt water. The released eDNA along with a tissue sample for each animal was then extracted and 16 single nucleotide polymorphisms (SNP) where genotyped.

Finally, **Chapter 6** provides a summary and synthesis of the findings of the empirical work. The broad applicability of how eDNA methods can be used to study species invasions and marine biodiversity are outlined along with further directions for study.

1.5 References

1992. Convention on Biological Diversity.*in* United_Nations, editor. C.N.29.1996, Rio de Janeiro.

Airoldi, L., X. Turon, S. Perkol-Finkel, and M. Rius. 2015. Corridors for aliens but not for natives: effects of marine urban sprawl at a regional scale. Diversity and Distributions **21**:755-768.

Andruszkiewicz, E. A., L. M. Sassoubre, and A. B. Boehm. 2017. Persistence of marine fish environmental DNA and the influence of sunlight. PLoS One **12**:e0185043.

Anton, A., N. R. Geraldi, C. E. Lovelock, E. T. Apostolaki, S. Bennett, J. Cebrian, D. Krause-Jensen, N. Marbà, P. Martinetto, J. M. Pandolfi, J. Santana-Garcon, and C. M. Duarte. 2019. Global ecological impacts of marine exotic species. Nature Ecology & Evolution **3**:787-800.

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

Ashton, G., I. Davidson, and G. Ruiz. 2014. Transient small boats as a long-distance coastal vector for dispersal of biofouling organisms. Estuaries and Coasts **37**:1572-1581.

Azzurro, E., F. Maynou, J. Belmaker, D. Golani, and J. A. Crooks. 2016. Lag times in Lessepsian fish invasion. Biological Invasions **18**:2761-2772.

Bailey, S. A. 2015. An overview of thirty years of research on ballast water as a vector for aquatic invasive species to freshwater and marine environments. Aquatic Ecosystem Health & Management **18**:261-268.

Baker, C. S., D. Steel, S. Nieukirk, and H. Klinck. 2018. Environmental DNA (eDNA) From the Wake of the Whales: Droplet Digital PCR for Detection and Species Identification. Frontiers in Marine Science **5**:133.

Bani, A., M. De Brauwer, S. Creer, A. J. Dumbrell, G. Limmon, J. Jompa, S. von der Heyden, and M. Beger. 2020. Chapter Ten - Informing marine spatial planning decisions with environmental DNA. Pages 375-407 *in* A. J. Dumbrell, E. C. Turner, and T. M. Fayle, editors. Advances in Ecological Research. Academic Press.

Barbier, E. B. 2017. Marine ecosystem services. Current Biology **27**:R507-R510.

Barnes, D. K. 2002. Invasions by marine life on plastic debris. Nature **416**:808-809.

Barnes, M. A., and C. R. Turner. 2016. The ecology of environmental DNA and implications for conservation genetics. Conservation Genetics **17**:1-17.

Barnett, T. P., D. W. Pierce, and R. Schnur. 2001. Detection of Anthropogenic Climate Change in the World's Oceans. Science **292**:270.

Bayer, S. R., P. D. Countway, and R. A. Wahle. 2019. Developing an eDNA toolkit to quantify broadcast spawning events of the sea scallop *Placopecten magellanicus*: moving beyond fertilization assays. Marine Ecology Progress Series **621**:127-141.

Beng, K. C., and R. T. Corlett. 2020. Applications of environmental DNA (eDNA) in ecology and conservation: opportunities, challenges and prospects. Biodiversity and Conservation **29**:2089-2121.

Beric, B., and H. J. MacIsaac. 2015. Determinants of rapid response success for alien invasive species in aquatic ecosystems. Biological Invasions **17**:3327-3335.

Bernatchez, L. 2019. Welcome to Environmental DNA! Environmental DNA **1**:3-4.

Bista, I., G. R. Carvalho, K. Walsh, M. Seymour, M. Hajibabaei, D. Lallias, M. Christmas, and S. Creer. 2017. Annual time-series analysis of aqueous eDNA reveals ecologically relevant dynamics of lake ecosystem biodiversity. Nature Communications **8**:14087.

Blackburn, T. M., P. Pyšek, S. Bacher, J. T. Carlton, R. P. Duncan, V. Jarošík, J. R. U. Wilson, and D. M. Richardson. 2011. A proposed unified framework for biological invasions. Trends in Ecology & Evolution **26**:333-339.

Blackman, R. C., K. K. S. Ling, L. R. Harper, P. Shum, B. Hanfling, and L. Lawson-Handley. 2020. Targeted and passive environmental DNA approaches outperform established methods for detection of quagga mussels, *Dreissena rostriformis bugensis* in flowing water. Ecology and Evolution **10**:13248-13259.

Blowes, S. A., S. R. Supp, L. H. Antao, A. Bates, H. Bruelheide, J. M. Chase, F. Moyes, A. Magurran, B. McGill, I. H. Myers-Smith, M. Winter, A. D. Bjorkman, D. E. Bowler, J. E. K. Byrnes, A. Gonzalez, J. Hines, F. Isbell, H. P.

Jones, L. M. Navarro, P. L. Thompson, M. Vellend, C. Waldock, and M. Dornelas. 2019. The geography of biodiversity change in marine and terrestrial assemblages. Science **366**:339-345.

Bohmann, K., A. Evans, M. T. P. Gilbert, G. R. Carvalho, S. Creer, M. Knapp, D. W. Yu, and M. de Bruyn. 2014. Environmental DNA for wildlife biology and biodiversity monitoring. Trends in Ecology & Evolution **29**:358-367.

Borrell, Y. J., L. Miralles, H. Do Huu, K. Mohammed-Geba, and E. Garcia-Vazquez. 2017. DNA in a bottle-rapid metabarcoding survey for early alerts of invasive species in ports. PLoS One **12**:e0183347.

Bourne, D. G., M. Garren, T. M. Work, E. Rosenberg, G. W. Smith, and C. D. Harvell. 2009. Microbial disease and the coral holobiont. Trends in Microbiology **17**:554-562.

Bourne, S. D., J. Hudson, L. E. Holman, and M. Rius. 2018. Marine Invasion Genomics: Revealing Ecological and Evolutionary Consequences of Biological Invasions.*in* O. MF and R. OM, editors. Population Genomics: Marine Organisms. Springer.

Boussarie, G., J. Bakker, O. S. Wangensteen, S. Mariani, L. Bonnin, J.-B. Juhel, J. J. Kiszka, M. Kulbicki, S. Manel, W. D. Robbins, L. Vigliola, and D. Mouillot. 2018. Environmental DNA illuminates the dark diversity of sharks. Science advances **4**:eaap9661.

Bowler, D. E., A. D. Bjorkman, M. Dornelas, I. H. Myers-Smith, L. M. Navarro, A. Niamir, S. R. Supp, C. Waldock, M. Winter, M. Vellend, S. A. Blowes, K. Böhning-Gaese, H. Bruelheide, R. Elahi, L. H. Antão, J. Hines, F. Isbell, H. P. Jones, A. E. Magurran, J. S. Cabral, and A. E. Bates. 2020. Mapping human pressures on biodiversity across the planet uncovers anthropogenic threat complexes. People and Nature **2**:380-394.

Brauko, K. M., A. Cabral, N. V. Costa, J. Hayden, C. E. Dias, E. S. Leite, R. D. Westphal, C. M. Mueller, J. M. Hall-Spencer, and R. R. Rodrigues. 2020. Marine heatwaves, sewage and eutrophication combine to trigger deoxygenation and biodiversity loss: a SW Atlantic case study. Frontiers in Marine Science **7**:**590258**.

Buckley, L. B., and W. Jetz. 2008. Linking global turnover of species and environments. Proceedings of the National Academy of Sciences **105**:17836.

Bugnot, A. B., M. Mayer-Pinto, L. Airoldi, E. C. Heery, E. L. Johnston, L. P. Critchley, E. M. A. Strain, R. L. Morris, L. H. L. Loke, M. J. Bishop, E. V. Sheehan, R. A. Coleman, and K. A. Dafforn. 2021. Current and projected global extent of marine built structures. Nature Sustainability **4**:33-41.

Buxton, A. S., J. J. Groombridge, and R. A. Griffiths. 2018. Seasonal variation in environmental DNA detection in sediment and water samples. PLoS One **13**:e0191737.

Caldwell, J., P. Payment, and R. Villemur. 2011. Mitochondrial DNA as source tracking markers of fecal contamination. Pages 229-250 Microbial source tracking: methods, applications, and case studies. Springer, New York.

Capinha, C., F. Essl, H. Seebens, D. Moser, and H. M. Pereira. 2015. The dispersal of alien species redefines biogeography in the Anthropocene. Science **348**:1248-1251.

Cardinale, B. J., J. E. Duffy, A. Gonzalez, D. U. Hooper, C. Perrings, P. Venail, A. Narwani, G. M. Mace, D. Tilman, D. A. Wardle, A. P. Kinzig, G. C. Daily, M. Loreau, J. B. Grace, A. Larigauderie, D. S. Srivastava, and S. Naeem. 2012. Biodiversity loss and its impact on humanity. Nature **486**:59-67.

Carim, K., N. Bean, J. Connor, W. Baker, M. Jaeger, M. Ruggles, K. McKelvey, T. Franklin, M. Young, and M. Schwartz. 2020. Environmental DNA sampling informs fish eradication efforts: case studies and lessons learned. North American Journal of Fisheries Management **40**:488-508.

Carlton, J. T., J. W. Chapman, J. B. Geller, J. A. Miller, D. A. Carlton, M. I. McCuller, N. C. Treneman, B. P. Steves, and G. M. Ruiz. 2017. Tsunami-driven rafting: Transoceanic species dispersal and implications for marine biogeography. Science **357**:1402-1406.

Carroll, E. L., M. W. Bruford, J. A. DeWoody, G. Leroy, A. Strand, L. Waits, and J. L. Wang. 2018. Genetic and genomic monitoring with minimally invasive sampling methods. Evolutionary Applications **11**:1094-1119.

Ceballos, G., P. R. Ehrlich, and R. Dirzo. 2017. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. Proceedings of the National Academy of Sciences **114**:E6089.

Chan, F. T., K. Stanislawczyk, A. C. Sneekes, A. Dvoretsky, S. Gollasch, D. Minchin, M. David, A. Jelmert, J. Albretsen, and S. A. Bailey. 2019. Climate change opens new frontiers for marine species in the Arctic: Current trends and future invasion risks. Global Change Biology **25**:25-38.

Chapman, A. D. 2009. Numbers of living species in Australia and the world.*in* Australian_Biodiversity_Information_Services, editor. Australian Goverment Department of Agriculture, Water and the Environment, Toowoomba, Australia.

Chen, J. S., E. Ma, L. B. Harrington, M. Da Costa, X. Tian, J. M. Palefsky, and J. A. Doudna. 2018. CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. Science **360**:436-439.

Clark, M. R., F. Althaus, T. A. Schlacher, A. Williams, D. A. Bowden, and A. A. Rowden. 2016. The impacts of deep-sea fisheries on benthic communities: a review. ICES Journal of Marine Science **73**:i51-i69.

Clusa, L., A. Ardura, S. Fernández, A. A. Roca, and E. García-Vázquez. 2017. An extremely sensitive nested PCR-RFLP mitochondrial marker for detection and identification of salmonids in eDNA from water samples. PeerJ **5**:e3045.

Colautti, R. I., and H. J. MacIsaac. 2004. A neutral terminology to define 'invasive' species. Diversity and Distributions **10**:135-141.

Coleman, F. C., and S. L. Williams. 2002. Overexploiting marine ecosystem engineers: potential consequences for biodiversity. Trends in Ecology & Evolution **17**:40-44.

Collins, R. A., O. S. Wangensteen, E. J. O'Gorman, S. Mariani, D. W. Sims, and M. J. Genner. 2018. Persistence of environmental DNA in marine systems. Communications Biology 1:185.

Costanza, R., R. d'Arge, R. de Groot, S. Farber, M. Grasso, B. Hannon, K. Limburg, S. Naeem, R. V. O'Neill, J. Paruelo, R. G. Raskin, P. Sutton, and M. van den Belt. 1997. The value of the world's ecosystem services and natural capital. Nature 387:253-260.

Costello, M. J., R. M. May, and N. E. Stork. 2013. Can We Name Earth's Species Before They Go Extinct? Science 339:413.

Costello, M. J., P. Tsai, P. S. Wong, A. K. L. Cheung, Z. Basher, and C. Chaudhary. 2017. Marine biogeographic realms and species endemicity. Nature Communications 8:1057.

Cowart, D. A., K. R. Murphy, and C. H. C. Cheng. 2018. Metagenomic sequencing of environmental DNA reveals marine faunal assemblages from the West Antarctic Peninsula. Marine Genomics 37:148-160.

Crain, C. M., K. Kroeker, and B. S. Halpern. 2008. Interactive and cumulative effects of multiple human stressors in marine systems. Ecology Letters 11:1304-1315.

Curtis, A. N., and E. R. Larson. 2020. No evidence that crayfish carcasses produce detectable environmental DNA (eDNA) in a stream enclosure experiment. PeerJ 8:e9333.

Darling, J. A., and J. T. Carlton. 2018. A Framework for Understanding Marine Cosmopolitanism in the Anthropocene. Frontiers in Marine Science 5:293.

Davidson, I. C., C. J. Zabin, A. L. Chang, C. W. Brown, M. D. Sytsma, and G. M. Ruiz. 2010. Recreational boats as potential vectors of marine organisms at an invasion hotspot. Aquatic Biology 11:179-191.

Davies, M. S., and S. Hawkins. 1998. Mucus from marine molluscs. Pages 1-71 Advances in marine biology. Elsevier.

Davison, P. I., G. H. Copp, V. Créach, L. Vilizzi, and J. Britton. 2017. Application of environmental DNA analysis to inform invasive fish eradication operations. The Science of Nature 104:35.

Deiner, K., and F. Altermatt. 2014. Transport Distance of Invertebrate Environmental DNA in a Natural River. PLoS One 9:e88786.

Deiner, K., H. M. Bik, E. Machler, M. Seymour, A. Lacoursiere-Roussel, F. Altermatt, S. Creer, I. Bista, D. M. Lodge, N. de Vere, M. E. Pfrender, and L. Bernatchez. 2017. Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. Molecular Ecology 26:5872-5895.

Deiner, K., E. A. Fronhofer, E. Mächler, J.-C. Walser, and F. Altermatt. 2016. Environmental DNA reveals that rivers are conveyer belts of biodiversity information. Nature Communications 7:12544.

Deiner, K., J. Lopez, S. Bourne, L. Holman, M. Seymour, E. K. Grey, A. Lacoursière, Y. Li, M. A. Renshaw, M. E. Pfrender, M. Rius, L. Bernatchez, and D. M. Lodge. 2018. Optimising the detection of marine taxonomic richness using environmental DNA metabarcoding: the effects of filter material, pore size and extraction method. Metabarcoding and Metagenomics 2:e28963.

Diagne, C., B. Leroy, A.-C. Vaissière, R. E. Gozlan, D. Roiz, I. Jarić, J.-M. Salles, C. J. A. Bradshaw, and F. Courchamp. 2021. High and rising economic costs of biological invasions worldwide. Nature 592:571-576.

Diederich, S., G. Nehls, J. E. Van Beusekom, and K. Reise. 2005. Introduced Pacific oysters (Crassostrea gigas) in the northern Wadden Sea: invasion accelerated by warm summers? Helgoland Marine Research 59:97-106.

Djurhuus, A., C. J. Closek, R. P. Kelly, K. J. Pitz, R. P. Michisaki, H. A. Starks, K. R. Walz, E. A. Andruszkiewicz, E. Olesin, K. Hubbard, E. Montes, D. Otis, F. E. Muller-Karger, F. P. Chavez, A. B. Boehm, and M. Breitbart. 2020. Environmental DNA reveals seasonal shifts and potential interactions in a marine community. Nature Communications 11:254.

Doi, H., K. Uchii, T. Takahara, S. Matsuhashi, H. Yamanaka, and T. Minamoto. 2015. Use of Droplet Digital PCR for Estimation of Fish Abundance and Biomass in Environmental DNA Surveys. PLoS One 10:e0122763.

Dornelas, M., N. J. Gotelli, B. McGill, H. Shimadzu, F. Moyes, C. Sievers, and A. E. Magurran. 2014. Assemblage time series reveal biodiversity change but not systematic loss. Science 344:296-299.

Dornelas, M., N. J. Gotelli, H. Shimadzu, F. Moyes, A. E. Magurran, and B. J. McGill. 2019a. A balance of winners and losers in the Anthropocene. Ecology Letters 22:847-854.

Dornelas, M., E. M. P. Madin, M. Bunce, J. D. DiBattista, M. Johnson, J. S. Madin, A. E. Magurran, B. J. McGill, N. Pettorelli, O. Pizarro, S. B. Williams, M. Winter, and A. E. Bates. 2019b. Towards a macroscope: Leveraging technology to transform the breadth, scale and resolution of macroecological data. Global Ecology and Biogeography 28:1937-1948.

Doyle, J., and S. Uthicke. 2020. Sensitive environmental DNA detection via lateral flow assay (dipstick)—A case study on corallivorous crown-of-thorns sea star (*Acanthaster cf. solaris*) detection. Environmental DNA 3.

Duarte, C. M., S. Agusti, E. Barbier, G. L. Britten, J. C. Castilla, J.-P. Gattuso, R. W. Fulweiler, T. P. Hughes, N. Knowlton, C. E. Lovelock, H. K. Lotze, M. Predragovic, E. Poloczanska, C. Roberts, and B. Worm. 2020. Rebuilding marine life. Nature 580:39-51.

Duarte, S., P. E. Vieira, A. S. Lavrador, and F. O. Costa. 2021. Status and prospects of marine NIS detection and monitoring through (e)DNA metabarcoding. Science of the Total Environment 751:141729.

Dufresnes, C., T. Déjean, S. Zumbach, B. R. Schmidt, L. Fumagalli, P. Ramseier, and S. Dubey. 2019. Early detection and spatial monitoring of an emerging biological invasion by population genetics and environmental DNA metabarcoding. Conservation Science and Practice 1:e86.

Dunker, K. J., A. J. Sepulveda, R. L. Massengill, J. B. Olsen, O. L. Russ, J. K. Wenburg, and A. Antonovich. 2016. Potential of Environmental DNA to Evaluate Northern Pike (*Esox lucius*) Eradication Efforts: An Experimental Test and Case Study. PLoS One **11**:e0162277.

Dunn, N., V. Priestley, A. Herraiz, R. Arnold, and V. Savolainen. 2017. Behavior and season affect crayfish detection and density inference using environmental DNA. Ecology and Evolution **7**:7777-7785.

Eichmiller, J. J., S. E. Best, and P. W. Sorensen. 2016. Effects of Temperature and Trophic State on Degradation of Environmental DNA in Lake Water. Environmental Science & Technology **50**:1859-1867.

Ekman, S. 1953. Zoogeography of the Sea. Sidgwick and Jackson.

Elahi, R., Mary I. O'Connor, Jarrett E. K. Byrnes, J. Dunic, Britas K. Eriksson, Marc J. S. Hensel, and Patrick J. Kearns. 2015. Recent Trends in Local-Scale Marine Biodiversity Reflect Community Structure and Human Impacts. Current Biology **25**:1938-1943.

Espinoza, G. J., J. M. Poland, and J. R. A. Bremer. 2017. Genotyping live fish larvae: Non-lethal and noninvasive DNA isolation from 3-5 day old hatchlings. Biotechniques **63**:181-186.

Essl, F., S. Bacher, P. Genovesi, P. E. Hulme, J. M. Jeschke, S. Katsanevakis, I. Kowarik, I. Kühn, P. Pyšek, W. Rabitsch, S. Schindler, M. van Kleunen, M. Vilà, J. R. U. Wilson, and D. M. Richardson. 2018. Which Taxa Are Alien? Criteria, Applications, and Uncertainties. BioScience **68**:496-509.

Essl, F., S. Dullinger, P. Genovesi, P. E. Hulme, J. M. Jeschke, S. Katsanevakis, I. Kühn, B. Lenzner, A. Pauchard, P. Pyšek, W. Rabitsch, D. M. Richardson, H. Seebens, M. van Kleunen, W. H. van der Putten, M. Vilà, and S. Bacher. 2019. A Conceptual Framework for Range-Expanding Species that Track Human-Induced Environmental Change. BioScience **69**:908-919.

EU. 2014. The prevention and management of the introduction and spread of invasive alien species.*in* E. Union, editor. 1143/2014.

Fernandes, J. A., L. Santos, T. Vance, T. Fileman, D. Smith, J. D. D. Bishop, F. Viard, A. M. Queirós, G. Merino, E. Buisman, and M. C. Austen. 2016. Costs and benefits to European shipping of ballast-water and hull-fouling treatment: Impacts of native and non-indigenous species. Marine Policy **64**:148-155.

Ficetola, G. F., F. Mazel, and W. Thuiller. 2017. Global determinants of zoogeographical boundaries. Nature Ecology & Evolution **1**:0089.

Ficetola, G. F., C. Miaud, F. Pompanon, and P. Taberlet. 2008. Species detection using environmental DNA from water samples. Biology Letters **4**:423-425.

Ficetola, G. F., J. Poulenard, P. Sabatier, E. Messager, L. Gielly, A. Leloup, D. Etienne, J. Bakke, E. Malet, B. Fanget, E. Støren, J.-L. Reyss, P. Taberlet, and F. Arnaud. 2018. DNA from lake sediments reveals long-term ecosystem changes after a biological invasion. Science advances **4**:eaar4292.

Forsström, T., and A. Vasemägi. 2016. Can environmental DNA (eDNA) be used for detection and monitoring of introduced crab species in the Baltic Sea? Marine Pollution Bulletin **109**:350-355.

Fossøy, F., H. Brandsegg, R. Sivertsgård, O. Pettersen, B. K. Sandercock, Ø. Solem, K. Hindar, and T. A. Mo. 2020. Monitoring presence and abundance of two gyrodactylid ectoparasites and their salmonid hosts using environmental DNA. Environmental DNA **2**:53-62.

Fremier, A. K., K. M. Strickler, J. Parzych, S. Powers, and C. S. Goldberg. 2019. Stream Transport and Retention of Environmental DNA Pulse Releases in Relation to Hydrogeomorphic Scaling Factors. Environmental Science & Technology **53**:6640-6649.

Frühe, L., T. Cordier, V. Dully, H.-W. Breiner, G. Lentendu, J. Pawlowski, C. Martins, T. A. Wilding, and T. Stoeck. 2020. Supervised machine learning is superior to indicator value inference in monitoring the environmental impacts of salmon aquaculture using eDNA metabarcodes. Molecular Ecology **In Press**.

Galil, B. S., F. Boero, M. L. Campbell, J. T. Carlton, E. Cook, S. Fraschetti, S. Gollasch, C. L. Hewitt, A. Jelmert, E. Macpherson, A. Marchini, C. McKenzie, D. Minchin, A. Occhipinti-Ambrogi, H. Ojaveer, S. Olenin, S. Piraino, and G. M. Ruiz. 2015. 'Double trouble': the expansion of the Suez Canal and marine bioinvasions in the Mediterranean Sea. Biological Invasions **17**:973-976.

Gallardo, B., M. Clavero, M. I. Sanchez, and M. Vila. 2016. Global ecological impacts of invasive species in aquatic ecosystems. Global Change Biology **22**:151-163.

Gamfeldt, L., J. S. Lefcheck, J. E. K. Byrnes, B. J. Cardinale, J. E. Duffy, and J. N. Griffin. 2015. Marine biodiversity and ecosystem functioning: what's known and what's next? Oikos **124**:252-265.

Geraldi, N. R., A. Anton, J. Santana-Garcon, S. Bennett, N. Marbà, C. E. Lovelock, E. T. Apostolaki, J. Cebrian, D. Krause-Jensen, P. Martinetto, J. M. Pandolfi, and C. M. Duarte. 2020. Ecological effects of non-native species in marine ecosystems relate to co-occurring anthropogenic pressures. Global Change Biology **26**:1248-1258.

Gerbilskii, N. L., and A. Petrunkevitch. 1955. Intraspecific Biological Groups of Acipenserines and Their Reproduction in the Lower Regions of Rivers with Regulated Flow. Systematic Zoology **4**:86-92.

Gerhard, W. A., and C. K. Gunsch. 2019. Metabarcoding and machine learning analysis of environmental DNA in ballast water arriving to hub ports. Environment International **124**:312-319.

Ghosal, R., J. J. Eichmiller, B. A. Witthuhn, and P. W. Sorensen. 2018. Attracting Common Carp to a bait site with food reveals strong positive relationships between fish density, feeding activity, environmental DNA, and sex pheromone release that could be used in invasive fish management. Ecology and Evolution **8**:6714-6727.

Giebner, H., K. Langen, S. J. Bourlat, S. Kukowka, C. Mayer, J. J. Astrin, B. Misof, and V. G. Fonseca. 2020. Comparing diversity levels in environmental samples: DNA sequence capture and metabarcoding approaches using 18S and COI genes. Molecular Ecology Resources **20**:1333-1345.

Giovos, I., P. Kleitou, D. Poursanidis, I. Batjakas, G. Bernardi, F. Crocetta, N. Doumpas, S. Kalogirou, T. E. Kampouris, I. Keramidas, J. Langeneck, M. Maximiadi, E. Mitsou, V.-O. Stoilas, F. Tiralongo, G. Romanidis-Kyriakidis, N.-J. Xentidis, A. Zenetos, and S. Katsanevakis. 2019. Citizen-science for monitoring marine invasions and stimulating public engagement: a case project from the eastern Mediterranean. Biological Invasions **21**:3707-3721.

Gold, Z., J. Sprague, D. J. Kushner, E. Zerecero, and P. H. Barber. 2021. eDNA metabarcoding as a biomonitoring tool for marine protected areas. PLoS One **16(2)**:e0238557.

Goldberg, C. S., C. R. Turner, K. Deiner, K. E. Klymus, P. F. Thomsen, M. A. Murphy, S. F. Spear, A. McKee, S. J. Oyler-McCance, R. S. Cornman, M. B. Laramie, A. R. Mahon, R. F. Lance, D. S. Pilliod, K. M. Strickler, L. P. Waits, A. K. Fremier, T. Takahara, J. E. Herder, P. Taberlet, and M. Gilbert. 2016. Critical considerations for the application of environmental DNA methods to detect aquatic species. Methods in Ecology and Evolution **7**:1299-1307.

Gonzalez, A., B. J. Cardinale, G. R. H. Allington, J. Byrnes, K. Arthur Endsley, D. G. Brown, D. U. Hooper, F. Isbell, M. I. O'Connor, and M. Loreau. 2016. Estimating local biodiversity change: a critique of papers claiming no net loss of local diversity. Ecology **97**:1949-1960.

Goodwin, S., J. D. McPherson, and W. R. McCombie. 2016. Coming of age: ten years of next-generation sequencing technologies. Nature Reviews Genetics **17**:333-351.

Gootenberg, J. S., O. O. Abudayyeh, M. J. Kellner, J. Joung, J. J. Collins, and F. Zhang. 2018. Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6. Science **360**:439-444.

Grey, E. K., L. Bernatchez, P. Cassey, K. Deiner, M. Deveney, K. L. Howland, A. Lacoursiere-Roussel, S. C. Y. Leong, Y. Li, B. Olds, M. E. Pfrender, T. A. A. Prowse, M. A. Renshaw, and D. M. Lodge. 2018. Effects of sampling effort on biodiversity patterns estimated from environmental DNA metabarcoding surveys. Scientific Reports **8**:8843.

Grosholz, E. D., R. E. Crafton, R. E. Fontana, J. R. Pasari, S. L. Williams, and C. J. Zabin. 2015. Aquaculture as a vector for marine invasions in California. Biological Invasions **17**:1471-1484.

Haeckel, E. 1904. Kunstformen der Natur. Leipzig und Wien Verlad des Bibliographiichen Initituts.

Haile, J., R. Holdaway, K. Oliver, M. Bunce, M. T. P. Gilbert, R. Nielsen, K. Munch, S. Y. W. Ho, B. Shapiro, and E. Willerslev. 2007. Ancient DNA Chronology within Sediment Deposits: Are Paleobiological Reconstructions Possible and Is DNA Leaching a Factor? Molecular Biology and Evolution **24**:982-989.

Harnik, P. G., H. K. Lotze, S. C. Anderson, Z. V. Finkel, S. Finnegan, D. R. Lindberg, L. H. Liow, R. Lockwood, C. R. McClain, J. L. McGuire, A. O'Dea, J. M. Pandolfi, C. Simpson, and D. P. Tittensor. 2012. Extinctions in ancient and modern seas. Trends in Ecology & Evolution **27**:608-617.

Harvell, C. D., K. Kim, J. M. Burkholder, R. R. Colwell, P. R. Epstein, D. J. Grimes, E. E. Hofmann, E. K. Lipp, A. D. M. E. Osterhaus, R. M. Overstreet, J. W. Porter, G. W. Smith, and G. R. Vasta. 1999. Emerging Marine Diseases--Climate Links and Anthropogenic Factors. Science **285**:1505.

He, X., T. F. Sutherland, J. Pawlowski, and C. L. Abbott. 2019. Responses of foraminifera communities to aquaculture-derived organic enrichment as revealed by environmental DNA metabarcoding. Molecular Ecology **28**:1138-1153.

Henley, W. F., P. J. Grobler, and R. J. Neves. 2006. Non-invasive method to obtain DNA from freshwater mussels (*Bivalvia: Unionidae*). Journal of Shellfish Research **25**:975-977.

Hewson, I., J. B. Button, B. M. Gudenkauf, B. Miner, A. L. Newton, J. K. Gaydos, J. Wynne, C. L. Groves, G. Hendler, and M. Murray. 2014. Densovirus associated with sea-star wasting disease and mass mortality. Proceedings of the National Academy of Sciences **111**:17278-17283.

Holman, L. E., Y. Chng, and M. Rius. 2021a. How does eDNA decay affect metabarcoding experiments? Environmental DNA **In Press**.

Holman, L. E., M. de Bruyn, S. Creer, G. Carvalho, J. Robidart, and M. Rius. 2019a. Detection of introduced and resident marine species using environmental DNA metabarcoding of sediment and water. Scientific Reports **9**:11559.

Holman, L. E., M. de Bruyn, S. Creer, G. Carvalho, J. Robidart, and M. Rius. 2021b. Animals, protists and bacteria share marine biogeographic patterns. Nature Ecology & Evolution **In Press**.

Holman, L. E., C. M. Hollenbeck, T. J. Ashton, and I. A. Johnston. 2019b. Demonstration of the Use of Environmental DNA for the Non-Invasive Genotyping of a Bivalve Mollusk, the European Flat Oyster (*Ostrea edulis*). Frontiers in Genetics **10**:1159.

Holt, B. G., J. P. Lessard, M. K. Borregaard, S. A. Fritz, M. B. Araujo, D. Dimitrov, P. H. Fabre, C. H. Graham, G. R. Graves, K. A. Jonsson, D. Nogues-Bravo, Z. Wang, R. J. Whittaker, J. Fjeldsa, and C. Rahbek. 2013. An update of Wallace's zoogeographic regions of the world. Science **339**:74-78.

Hotchkiss, R. S., A. Strasser, J. E. McDunn, and P. E. Swanson. 2009. Cell Death. New England Journal of Medicine **361**:1570-1583.

Hulme, P. E., S. Bacher, M. Kenis, S. Klotz, I. Kühn, D. Minchin, W. Nentwig, S. Olenin, V. Panov, and J. Pergl. 2008. Grasping at the routes of biological invasions: a framework for integrating pathways into policy. Journal of Applied Ecology **45**:403-414.

Jackson, J. B. C., M. X. Kirby, W. H. Berger, K. A. Bjorndal, L. W. Botsford, B. J. Bourque, R. H. Bradbury, R. Cooke, J. Erlandson, J. A. Estes, T. P. Hughes, S. Kidwell, C. B. Lange, H. S. Lenihan, J. M. Pandolfi, C. H. Peterson, R. S. Steneck, M. J. Tegner, and R. R. Warner. 2001. Historical Overfishing and the Recent Collapse of Coastal Ecosystems. Science **293**:629.

Jarić, I., T. Heger, F. Castro Monzon, J. M. Jeschke, I. Kowarik, K. R. McConkey, P. Pyšek, A. Sagouis, and F. Essl. 2019. Crypticity in Biological Invasions. Trends in Ecology & Evolution **34**:291-302.

Jarman, S. N., O. Berry, and M. Bunce. 2018. The value of environmental DNA biobanking for long-term biomonitoring. Nature Ecology and Evolution **2**:1192-1193.

Jensen, Ø., T. Dempster, E. Thorstad, I. Uglem, and A. Fredheim. 2010. Escapes of fishes from Norwegian sea-cage aquaculture: causes, consequences and prevention. Aquaculture Environment Interactions **1**:71-83.

Jerde, C. L. 2019. Can we manage fisheries with the inherent uncertainty from eDNA? Journal of Fish Biology **98**.

Jerde, C. L., B. P. Olds, A. J. Shogren, E. A. Andruszkiewicz, A. R. Mahon, D. Bolster, and J. L. Tank. 2016. Influence of Stream Bottom Substrate on Retention and Transport of Vertebrate Environmental DNA. Environmental Science & Technology **50**:8770-8779.

Jo, T., M. Arimoto, H. Murakami, R. Masuda, and T. Minamoto. 2019a. Particle size distribution of environmental DNA from the nuclei of marine fish. Environmental Science & Technology **53**:9947-9956.

Jo, T., H. Murakami, S. Yamamoto, R. Masuda, and T. Minamoto. 2019b. Effect of water temperature and fish biomass on environmental DNA shedding, degradation, and size distribution. Ecology and Evolution **9**:1135-1146.

Kamoroff, C., and C. S. Goldberg. 2018. An issue of life or death: using eDNA to detect viable individuals in wilderness restoration. Freshwater Science **37**:685-696.

Katsanevakis, S., A. Zenetos, C. Belchior, and A. C. Cardoso. 2013. Invading European Seas: Assessing pathways of introduction of marine aliens. Ocean & Coastal Management **76**:64-74.

Kelly, R. P., A. O. Shelton, and R. Gallego. 2019. Understanding PCR Processes to Draw Meaningful Conclusions from Environmental DNA Studies. Scientific Reports **9**:12133.

Klymus, K. E., C. A. Richter, D. C. Chapman, and C. Paukert. 2015. Quantification of eDNA shedding rates from invasive bighead carp *Hypophthalmichthys nobilis* and silver carp *Hypophthalmichthys molitrix*. Biological Conservation **183**:77-84.

Knudsen, S. W., R. B. Ebert, M. Hesselsøe, F. Kuntke, J. Hassingboe, P. B. Mortensen, P. F. Thomsen, E. E. Sigsgaard, B. K. Hansen, E. E. Nielsen, and P. R. Møller. 2019. Species-specific detection and quantification of environmental DNA from marine fishes in the Baltic Sea. Journal of Experimental Marine Biology and Ecology **510**:31-45.

Kocsis, Á. T., C. J. Reddin, and W. Kiessling. 2018. The stability of coastal benthic biogeography over the last 10 million years. Global Ecology and Biogeography **27**:1106-1120.

Koziol, A., M. Stat, T. Simpson, S. Jarman, J. D. DiBattista, E. S. Harvey, M. Marnane, J. McDonald, and M. Bunce. 2019. Environmental DNA metabarcoding studies are critically affected by substrate selection. Molecular Ecology Resources **19**:366-376.

Lacoursière-Roussel, A., M. Rosabal, and L. Bernatchez. 2016. Estimating fish abundance and biomass from eDNA concentrations: variability among capture methods and environmental conditions. Molecular Ecology Resources **16**:1401-1414.

Le Vin, A. L., A. Adam, A. Tedder, K. E. Arnold, and B. K. Mable. 2011. Validation of swabs as a non-destructive and relatively non-invasive DNA sampling method in fish. Molecular Ecology Resources **11**:107-109.

Leignel, V., J. H. Stillman, S. Baringou, R. Thabet, and I. Metais. 2014. Overview on the European green crab Carcinus spp. (*Portunidae, Decapoda*), one of the most famous marine invaders and ecotoxicological models. Environmental Science and Pollution Research **21**:9129-9144.

Leray, M., J. Y. Yang, C. P. Meyer, S. C. Mills, N. Agudelo, V. Ranwez, J. T. Boehm, and R. J. Machida. 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.

Leung, B., D. M. Lodge, D. Finnoff, J. F. Shogren, M. A. Lewis, and G. Lamberti. 2002. An ounce of prevention or a pound of cure: bioeconomic risk analysis of invasive species. Proceedings of the Royal Society of London. Series B: Biological Sciences **269**:2407-2413.

Li, J., L. J. Lawson Handley, L. R. Harper, R. Brys, H. V. Watson, C. Di Muri, X. Zhang, and B. Hänfling. 2019. Limited dispersion and quick degradation of environmental DNA in fish ponds inferred by metabarcoding. Environmental DNA **1**:238-250.

Lindahl, T. 1993. Instability and decay of the primary structure of DNA. Nature **362**:709-715.

Liu, X., T. M. Blackburn, T. Song, X. Li, C. Huang, and Y. Li. 2019. Risks of biological invasion on the belt and road. Current Biology **29**:499-505. e494.

Locey, K. J., and J. T. Lennon. 2016. Scaling laws predict global microbial diversity. Proceedings of the National Academy of Sciences **113**:5970.

Lorenz, M. G., and W. Wackernagel. 1987. Adsorption of DNA to sand and variable degradation rates of adsorbed DNA. Appl Environ Microbiol **53**:2948-2952.

Mächler, E., M. Osathanunkul, and F. Altermatt. 2018. Shedding light on eDNA: neither natural levels of UV radiation nor the presence of a filter feeder affect eDNA-based detection of aquatic organisms. PLoS One **13**:e0195529.

Magellan, K. 2019. Prayer animal release: An understudied pathway for introduction of invasive aquatic species. Aquatic Ecosystem Health & Management **22**:452-461.

Magris, R. A., and N. C. Ban. 2019. A meta-analysis reveals global patterns of sediment effects on marine biodiversity. Global Ecology and Biogeography **28**:1879-1898.

Magurran, A. E. 2013. Measuring biological diversity. John Wiley & Sons.

Marbuah, G., I.-M. Gren, and B. McKie. 2014. Economics of harmful invasive species: a review. Diversity **6**:500-523.

Martellini, A., P. Payment, and R. Villemur. 2005. Use of eukaryotic mitochondrial DNA to differentiate human, bovine, porcine and ovine sources in fecally contaminated surface water. Water Research **39**:541-548.

Maruyama, A., K. Nakamura, H. Yamanaka, M. Kondoh, and T. Minamoto. 2014. The Release Rate of Environmental DNA from Juvenile and Adult Fish. PLoS One **9**:e114639.

Mauvisseau, Q., J. Davy-Bowker, M. Bulling, R. Brys, S. Neyrinck, C. Troth, and M. Sweet. 2019. Combining ddPCR and environmental DNA to improve detection capabilities of a critically endangered freshwater invertebrate. Scientific Reports **9**:14064.

Mazor, T., C. Doropoulos, F. Schwarzmueller, D. W. Gladish, N. Kumaran, K. Merkel, M. Di Marco, and V. Gagic. 2018. Global mismatch of policy and research on drivers of biodiversity loss. Nature Ecology & Evolution **2**:1071-1074.

McCarthy, A. H., L. S. Peck, K. A. Hughes, and D. C. Aldridge. 2019. Antarctica: The final frontier for marine biological invasions. Global Change Biology **25**:2221-2241.

McClenachan, L., G. O'Connor, and T. Reynolds. 2015. Adaptive capacity of co-management systems in the face of environmental change: The soft-shell clam fishery and invasive green crabs in Maine. Marine Policy **52**:26-32.

McElwain, J. C., and S. W. Punyasena. 2007. Mass extinction events and the plant fossil record. Trends in Ecology & Evolution **22**:548-557.

McGill, B. J., M. Dornelas, N. J. Gotelli, and A. E. Magurran. 2015. Fifteen forms of biodiversity trend in the Anthropocene. Trends in Ecology & Evolution **30**:104-113.

Merkes, C. M., S. G. McCalla, N. R. Jensen, M. P. Gaikowski, and J. J. Amberg. 2014. Persistence of DNA in Carcasses, Slime and Avian Feces May Affect Interpretation of Environmental DNA Data. PLoS One **9**:e113346.

Millennium_Ecosystem_Assessment. 2005. Ecosystems and Human Well-being: Biodiversity Synthesis. World Resources Institute, Washington, DC.

Miller, A. W., and G. M. Ruiz. 2014. Arctic shipping and marine invaders. Nature Climate Change **4**:413-416.

Minamoto, T., M. Fukuda, K. R. Katsuhara, A. Fujiwara, S. Hidaka, S. Yamamoto, K. Takahashi, and R. Masuda. 2017. Environmental DNA reflects spatial and temporal jellyfish distribution. PLoS One **12**:e0173073.

Minamoto, T., M. Miya, T. Sado, S. Seino, H. Doi, M. Kondoh, K. Nakamura, T. Takahara, S. Yamamoto, H. Yamanaka, H. Araki, W. Iwasaki, A. Kasai, R. Masuda, and K. Uchii. 2020. An illustrated manual for environmental DNA research: Water sampling guidelines and experimental protocols. Environmental DNA **3**:8-13.

Miralles, L., E. Dopico, F. Devlo-Delva, and E. Garcia-Vazquez. 2016. Controlling populations of invasive pygmy mussel (*Xenostrobus securis*) through citizen science and environmental DNA. Marine Pollution Bulletin **110**:127-132.

Miya, M., R. O. Gotoh, and T. Sado. 2020. MiFish metabarcoding: a high-throughput approach for simultaneous detection of multiple fish species from environmental DNA and other samples. Fisheries Science **86**:939–970.

Molnar, J. L., R. L. Gamboa, C. Revenga, and M. D. Spalding. 2008. Assessing the global threat of invasive species to marine biodiversity. Frontiers in Ecology and the Environment **6**:485-492.

Moser, C. S., T. P. Wier, M. R. First, J. F. Grant, S. C. Riley, S. H. Robbins-Wamsley, M. N. Tamburri, G. M. Ruiz, A. W. Miller, and L. A. Drake. 2017. Quantifying the extent of niche areas in the global fleet of commercial ships: the potential for "super-hot spots" of biofouling. Biological Invasions **19**:1745-1759.

Moushomi, R., G. Wilgar, G. Carvalho, S. Creer, and M. Seymour. 2019. Environmental DNA size sorting and degradation experiment indicates the state of *Daphnia magna* mitochondrial and nuclear eDNA is subcellular. Scientific Reports **9**:12500.

Muirhead, J. R., M. S. Minton, W. A. Miller, and G. M. Ruiz. 2015. Projected effects of the Panama Canal expansion on shipping traffic and biological invasions. Diversity and Distributions **21**:75-87.

Murakami, H., S. Yoon, A. Kasai, T. Minamoto, S. Yamamoto, M. K. Sakata, T. Horiuchi, H. Sawada, M. Kondoh, Y. Yamashita, and R. Masuda. 2019. Dispersion and degradation of environmental DNA from caged fish in a marine environment. Fisheries Science **85**:327-337.

Murchie, T. J., M. Kuch, A. T. Duggan, M. L. Ledger, K. Roche, J. Klunk, E. Karpinski, D. Hackenberger, T. Sadoway, R. MacPhee, D. Froese, and H. Poinar. 2020. Optimizing extraction and targeted capture of ancient environmental DNA for reconstructing past environments using the PalaeoChip Arctic-1.0 bait-set. Quaternary Research **99**:305-328.

Nascimento, F. J. A., M. Dahl, D. Deyanova, L. D. Lyimo, H. M. Bik, T. Schuelke, T. J. Pereira, M. Björk, S. Creer, and M. Gullström. 2019. Above-below surface interactions mediate effects of seagrass disturbance on meiobenthic diversity, nematode and polychaete trophic structure. Communications Biology **2**:362.

Naviaux, R. K., B. Good, J. D. McPherson, D. L. Steffen, D. Markusic, B. Ransom, and J. Corbeil. 2005. Sand DNA—a genetic library of life at the water's edge. Marine Ecology Progress Series **301**:9-22.

Nelson-Chorney, H. T., C. S. Davis, M. S. Poesch, R. D. Vinebrooke, C. M. Carli, and M. K. Taylor. 2019. Environmental DNA in lake sediment reveals biogeography of native genetic diversity. Frontiers in Ecology and the Environment **17**:313– 318.

Padilla, D. K., and S. L. Williams. 2004. Beyond ballast water: aquarium and ornamental trades as sources of invasive species in aquatic ecosystems. Frontiers in Ecology and the Environment 2:131-138.

Pandolfi, J. M., T. L. Staples, and W. Kiessling. 2020. Increased extinction in the emergence of novel ecological communities. Science 370:220.

Paul, J. H., W. H. Jeffrey, and M. F. DeFlaun. 1987. Dynamics of extracellular DNA in the marine environment. Appl Environ Microbiol 53:170-179.

Pawlowski, J., P. Esling, F. Lejzerowicz, T. Cedhagen, and T. A. Wilding. 2014. Environmental monitoring through protist next-generation sequencing metabarcoding: assessing the impact of fish farming on benthic foraminifera communities. Molecular Ecology Resources 14:1129-1140.

Pecl, G. T., M. B. Araujo, J. D. Bell, J. Blanchard, T. C. Bonebrake, I. C. Chen, T. D. Clark, R. K. Colwell, F. Danielsen, B. Evengard, L. Falconi, S. Ferrier, S. Frusher, R. A. Garcia, R. B. Griffis, A. J. Hobday, C. Janion-Scheepers, M. A. Jarzyna, S. Jennings, J. Lenoir, H. I. Linnetved, V. Y. Martin, P. C. McCormack, J. McDonald, N. J. Mitchell, T. Mustonen, J. M. Pandolfi, N. Pettorelli, E. Popova, S. A. Robinson, B. R. Scheffers, J. D. Shaw, C. J. Sorte, J. M. Strugnell, J. M. Sunday, M. N. Tuanmu, A. Verges, C. Villanueva, T. Wernberg, E. Wapstra, and S. E. Williams. 2017. Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. Science 355:6332.

Pedersen, M. W., A. Ginolhac, L. Orlando, J. Olsen, K. Andersen, J. Holm, S. Funder, E. Willerslev, and K. H. Kjær. 2013. A comparative study of ancient environmental DNA to pollen and macrofossils from lake sediments reveals taxonomic overlap and additional plant taxa. Quaternary Science Reviews 75:161-168.

Pedersen, M. W., S. Overballe-Petersen, L. Ermini, C. Der Sarkissian, J. Haile, M. Hellstrom, J. Spens, P. F. Thomsen, K. Bohmann, and E. Cappellini. 2015. Ancient and modern environmental DNA. Philosophical Transactions of the Royal Society B: Biological Sciences 370:20130383.

Pedersen, M. W., A. Ruter, C. Schweger, H. Friebe, R. A. Staff, K. K. Kjeldsen, M. L. Mendoza, A. B. Beaudoin, C. Zutter, and N. K. Larsen. 2016. Postglacial viability and colonization in North America's ice-free corridor. Nature 537:45.

Peters, L., S. Spatharis, M. A. Dario, T. Dwyer, I. J. T. Roca, A. Kintner, O. Kanstad-Hanssen, M. S. Llewellyn, and K. Praebel. 2018. Environmental DNA: A New Low-Cost Monitoring Tool for Pathogens in Salmonid Aquaculture. Frontiers in Microbiology 9:3009.

Pietramellara, G., J. Ascher, F. Borgogni, M. T. Ceccherini, G. Guerri, and P. Nannipieri. 2009. Extracellular DNA in soil and sediment: fate and ecological relevance. Biology and Fertility of Soils 45:219-235.

Pietramellara, G., J. Ascher, M. T. Ceccherini, P. Nannipieri, and D. Wenderoth. 2007. Adsorption of pure and dirty bacterial DNA on clay minerals and their transformation frequency. Biology and Fertility of Soils 43:731-739.

Pimm, S. L., C. N. Jenkins, R. Abell, T. M. Brooks, J. L. Gittleman, L. N. Joppa, P. H. Raven, C. M. Roberts, and J. O. Sexton. 2014. The biodiversity of species and their rates of extinction, distribution, and protection. Science 344:6187.

Pineda, M. C., S. López-Legentil, and X. Turon. 2011. The Whereabouts of an Ancient Wanderer: Global Phylogeography of the Solitary Ascidian Styela plicata. PLoS One 6:e25495.

Plough, L. V., M. B. Ogburn, C. L. Fitzgerald, R. Geranio, G. A. Marafino, and K. D. Richie. 2018. Environmental DNA analysis of river herring in Chesapeake Bay: A powerful tool for monitoring threatened keystone species. PLoS One 13:e0205578.

Polidoro, B. A., K. E. Carpenter, L. Collins, N. C. Duke, A. M. Ellison, J. C. Ellison, E. J. Farnsworth, E. S. Fernando, K. Kathiresan, N. E. Koedam, S. R. Livingstone, T. Miyagi, G. E. Moore, V. Ngoc Nam, J. E. Ong, J. H. Primavera, S. G. Salmo, III, J. C. Sanciangco, S. Sukardjo, Y. Wang, and J. W. H. Yong. 2010. The Loss of Species: Mangrove Extinction Risk and Geographic Areas of Global Concern. PLoS One 5:e10095.

Postaire, B. D., J. Bakker, J. Gardiner, T. R. Wiley, and D. D. Chapman. 2020. Environmental DNA detection tracks established seasonal occurrence of blacktip sharks (*Carcharhinus limbatus*) in a semi-enclosed subtropical bay. Scientific Reports 10:11847.

Proches, S., and S. Ramdhani. 2012. The World's Zoogeographical Regions Confirmed by Cross-Taxon Analyses. BioScience 62:260-270.

Pyšek, P., and D. M. Richardson. 2010. Invasive Species, Environmental Change and Management, and Health. Annual Review of Environment and Resources 35:25-55.

Pyšek, P., D. M. Richardson, M. Rejmánek, G. L. Webster, M. Williamson, and J. Kirschner. 2004. Alien plants in checklists and floras: towards better communication between taxonomists and ecologists. TAXON 53:131-143.

Quince, C., A. W. Walker, J. T. Simpson, N. J. Loman, and N. Segata. 2017. Shotgun metagenomics, from sampling to analysis. Nature Biotechnology 35:833-844.

Ratcliffe, F. C., T. M. Uren Webster, C. Garcia de Leaniz, and S. Consuegra. 2021. A drop in the ocean: Monitoring fish communities in spawning areas using environmental DNA. Environmental DNA 3:43-54.

Rey, A., K. J. Carney, L. E. Quinones, K. M. Pagenkopp Lohan, G. M. Ruiz, O. C. Basurko, and N. Rodríguez-Ezpeleta. 2019. Environmental DNA Metabarcoding: A Promising Tool for Ballast Water Monitoring. Environmental Science & Technology 53:11849-11859.

Ricciardi, A., and R. Kipp. 2008. Predicting the number of ecologically harmful exotic species in an aquatic system. Diversity and Distributions 14:374-380.

Richardson, D. M., P. Pyšek, M. Rejmánek, M. G. Barbour, F. D. Panetta, and C. J. West. 2000. Naturalization and invasion of alien plants: concepts and definitions. Diversity and Distributions **6**:93-107.

Richter, D. J., R. Watteaux, T. Vannier, J. Leconte, P. Frémont, G. Reygondeau, N. Maillet, N. Henry, G. Benoit, A. Fernàndez-Guerra, S. Suweis, R. Narci, C. Berney, D. Eveillard, F. Gavory, L. Guidi, K. Labadie, E. Mahieu, J. Poulain, S. Romac, S. Roux, C. Dimier, S. Kandels, M. Picheral, S. Searson, S. Pesant, J.-M. Aury, J. R. Brum, C. Lemaitre, E. Pelletier, P. Bork, S. Sunagawa, L. Karp-Boss, C. Bowler, M. B. Sullivan, E. Karsenti, M. Mariadassou, I. Probert, P. Peterlongo, P. Wincker, C. de Vargas, M. Ribera d'Alcalà, D. Iudicone, and O. Jaillon. 2019. Genomic evidence for global ocean plankton biogeography shaped by large-scale current systems. bioRxiv:867739.

Robson, H. L. A., T. H. Noble, R. J. Saunders, S. K. A. Robson, D. W. Burrows, and D. R. Jerry. 2016. Fine-tuning for the tropics: application of eDNA technology for invasive fish detection in tropical freshwater ecosystems. Molecular Ecology Resources **16**:922-932.

Rodriguez-Ezpeleta, N., O. Morissette, C. Bean, S. Manu, P. Banerjee, A. Lacoursiere, K. Ben, E. Alter, F. Roger, and L. Holman. 2020. Trade-offs between reducing complex terminology and producing accurate interpretations from environmental DNA: Comment on "Environmental DNA: What's behind the term?" by Pawlowski et al.(2020). Authorea Preprints.

Rodriguez-Ezpeleta, N., O. Morissette, C. W. Bean, S. Manu, P. Banerjee, A. Lacoursière-Roussel, K. C. Beng, S. E. Alter, F. Roger, L. E. Holman, K. A. Stewart, M. T. Monaghan, Q. Mauvisseau, L. Mirimin, O. S. Wangensteen, C. M. Antognazza, S. J. Helyar, H. de Boer, M.-E. Monchamp, R. Nijland, C. L. Abbott, H. Doi, M. A. Barnes, M. Leray, P. I. Hablützel, and K. Deiner. 2021. Trade-offs between reducing complex terminology and producing accurate interpretations from environmental DNA: Comment on "Environmental DNA: What's behind the term?" by Pawlowski et al., (2020). Molecular Ecology **In Press**.

Rossolini, G. M., S. Cresti, A. Ingianni, P. Cattani, M. L. Riccio, and G. Satta. 1994. Use of deoxyinosine-containing primers vs degenerate primers for polymerase chain reaction based on ambiguous sequence information. Molecular and Cellular Probes **8**:91-98.

Ruiz, G. M., P. W. Fofonoff, B. P. Steves, and J. T. Carlton. 2015. Invasion history and vector dynamics in coastal marine ecosystems: A North American perspective. Aquatic Ecosystem Health & Management **18**:299-311.

Sakata, M. K., S. Yamamoto, R. O. Gotoh, M. Miya, H. Yamanaka, and T. Minamoto. 2020. Sedimentary eDNA provides different information on timescale and fish species composition compared with aqueous eDNA. Environmental DNA **2**:505-518.

Salter, I. 2018. Seasonal variability in the persistence of dissolved environmental DNA (eDNA) in a marine system: The role of microbial nutrient limitation. PLoS One **13**:e0192409.

Salter, I., M. Joensen, R. Kristiansen, P. Steingrund, and P. Vestergaard. 2019. Environmental DNA concentrations are correlated with regional biomass of Atlantic cod in oceanic waters. Communications Biology **2**:461.

Sancar, A., and G. B. Sancar. 1988. DNA Repair Enzymes. Annual Review of Biochemistry **57**:29-67.

Sardain, A., E. Sardain, and B. Leung. 2019. Global forecasts of shipping traffic and biological invasions to 2050. Nature Sustainability **2**:274-282.

Sassoubre, L. M., K. M. Yamahara, L. D. Gardner, B. A. Block, and A. B. Boehm. 2016. Quantification of environmental DNA (eDNA) shedding and decay rates for three marine fish. Environmental Science & Technology **50**:10456-10464.

Schloss, P. D., R. A. Girard, T. Martin, J. Edwards, and J. C. Thrash. 2016. Status of the Archaeal and Bacterial Census: an Update. MBio **7**:e00201-00216.

Schnell, I. B., K. Bohmann, S. E. Schultze, S. R. Richter, D. C. Murray, M.-H. S. Sinding, D. Bass, J. E. Cadle, M. J. Campbell, R. Dolch, D. P. Edwards, T. N. E. Gray, T. Hansen, A. N. Q. Hoa, C. L. Noer, S. Heise-Pavlov, A. F. Sander Pedersen, J. C. Ramamonjisoa, M. E. Siddall, A. Tilker, C. Traeholt, N. Wilkinson, P. Woodcock, D. W. Yu, M. F. Bertelsen, M. Bunce, and M. T. P. Gilbert. 2018. Debugging diversity – a pan-continental exploration of the potential of terrestrial blood-feeding leeches as a vertebrate monitoring tool. Molecular Ecology Resources **18**:1282-1298.

Scholes, R., G. Mace, W. Turner, G. Geller, N. Jürgens, A. Larigauderie, D. Muchoney, B. Walther, and H. Mooney. 2008. Toward a global biodiversity observing system. Science **321**:1044-1045.

Schroeder, G. K., and R. Wolfenden. 2007. Rates of Spontaneous Disintegration of DNA and the Rate Enhancements Produced by DNA Glycosylases and Deaminases. Biochemistry **46**:13638-13647.

Scyphers, S. B., S. P. Powers, J. L. Akins, J. M. Drymon, C. W. Martin, Z. H. Schobernd, P. J. Schofield, R. L. Shipp, and T. S. Switzer. 2015. The Role of Citizens in Detecting and Responding to a Rapid Marine Invasion. Conservation Letters **8**:242-250.

Secretariat_of_the_Convention_on_Biological_Diversity. 2020. Global Biodiversity Outlook 5. Montreal.

Seymour, M., I. Durance, B. J. Cosby, E. Ransom-Jones, K. Deiner, S. J. Ormerod, J. K. Colbourne, G. Wilgar, G. R. Carvalho, M. de Bruyn, F. Edwards, B. A. Emmett, H. M. Bik, and S. Creer. 2018. Acidity promotes degradation of multi-species environmental DNA in lotic mesocosms. Communications Biology **1**:4.

Shephard, K. L. 1994. Functions for fish mucus. Reviews in fish biology and fisheries **4**:401-429.

Shogren, A. J., J. L. Tank, E. Andruszkiewicz, B. Olds, A. R. Mahon, C. L. Jerde, and D. Bolster. 2017. Controls on eDNA movement in streams: Transport, Retention, and Resuspension. Scientific Reports **7**:5065.

Shogren, A. J., J. L. Tank, E. A. Andruszkiewicz, B. Olds, C. Jerde, and D. Bolster. 2016. Modelling the transport of environmental DNA through a porous substrate using continuous flow-through column experiments. Journal of The Royal Society Interface **13**:20160290.

Signor, P. W. 1994. Biodiversity in Geological Time. American Zoologist **34**:23-32.

Sigsgaard, E. E., I. B. Nielsen, H. Carl, M. A. Krag, S. W. Knudsen, Y. Xing, T. H. Holm-Hansen, P. R. Møller, and P. F. Thomsen. 2017. Seawater environmental DNA reflects seasonality of a coastal fish community. Marine Biology **164**:128.

Simberloff, D., J.-L. Martin, P. Genovesi, V. Maris, D. A. Wardle, J. Aronson, F. Courchamp, B. Galil, E. García-Berthou, M. Pascal, P. Pyšek, R. Sousa, E. Tabacchi, and M. Vilà. 2013. Impacts of biological invasions: what's what and the way forward. Trends in Ecology & Evolution **28**:58-66.

Spalding, M. D., H. E. Fox, G. R. Allen, N. Davidson, Z. A. Ferdaña, M. Finlayson, B. S. Halpern, M. A. Jorge, A. Lombana, S. A. Lourie, K. D. Martin, E. McManus, J. Molnar, C. A. Recchia, and J. Robertson. 2007. Marine Ecoregions of the World: A Bioregionalization of Coastal and Shelf Areas. BioScience **57**:573-583.

Stat, M., M. J. Huggett, R. Bernasconi, J. D. DiBattista, T. E. Berry, S. J. Newman, E. S. Harvey, and M. Bunce. 2017. Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. Scientific Reports **7**:12240.

Stewart, K. A. 2019. Understanding the effects of biotic and abiotic factors on sources of aquatic environmental DNA. Biodiversity and Conservation **28**:983-1001.

Stoeckle, M. Y., L. Soboleva, and Z. Charlop-Powers. 2017. Aquatic environmental DNA detects seasonal fish abundance and habitat preference in an urban estuary. PLoS One **12**:e0175186.

Sunagawa, S., L. P. Coelho, S. Chaffron, J. R. Kultima, K. Labadie, G. Salazar, B. Djahanschiri, G. Zeller, D. R. Mende, A. Alberti, F. M. Cornejo-Castillo, P. I. Costea, C. Cruaud, F. d'Ovidio, S. Engelen, I. Ferrera, J. M. Gasol, L. Guidi, F. Hildebrand, F. Kokoszka, C. Lepoivre, G. Lima-Mendez, J. Poulain, B. T. Poulos, M. Royo-Llonch, H. Sarmento, S. Vieira-Silva, C. Dimier, M. Picheral, S. Searson, S. Kandels-Lewis, C. Bowler, C. de Vargas, G. Gorsky, N. Grimsley, P. Hingamp, D. Iudicone, O. Jaillon, F. Not, H. Ogata, S. Pesant, S. Speich, L. Stemmann, M. B. Sullivan, J. Weissenbach, P. Wincker, E. Karsenti, J. Raes, S. G. Acinas, and P. Bork. 2015. Structure and function of the global ocean microbiome. Science **348**:1261359.

Sunday, J. M., A. E. Bates, and N. K. Dulvy. 2012. Thermal tolerance and the global redistribution of animals. Nature Climate Change **2**:686-690.

Taberlet, P., A. Bonin, L. Zinger, and E. Coissac. 2018. Environmental DNA : for biodiversity research and monitoring. Oxford University Press, Oxford, United Kingdom.

Takahara, T., T. Minamoto, H. Yamanaka, H. Doi, and Z. i. Kawabata. 2012. Estimation of Fish Biomass Using Environmental DNA. PLoS One **7**:e35868.

Taslima, K., A. Davie, B. J. McAndrew, and D. J. Penman. 2016. DNA sampling from mucus in the Nile tilapia, *Oreochromis niloticus*: minimally invasive sampling for aquaculture-related genetics research. Aquaculture Research **47**:4032-4037.

Thalinger, B., K. Deiner, L. R. Harper, H. C. Rees, R. C. Blackman, D. Sint, M. Traugott, C. S. Goldberg, and K. Bruce. 2021. A validation scale to determine the readiness of environmental DNA assays for routine species monitoring. Environmental DNA **In Press**.

Theobald, E. J., A. K. Ettinger, H. K. Burgess, L. B. DeBey, N. R. Schmidt, H. E. Froehlich, C. Wagner, J. HilleRisLambers, J. Tewksbury, M. A. Harsch, and J. K. Parrish. 2015. Global change and local solutions: Tapping the unrealized potential of citizen science for biodiversity research. Biological Conservation **181**:236-244.

Thomsen, M. S., J. E. Byers, D. R. Schiel, J. F. Bruno, J. D. Olden, T. Wernberg, and B. R. Silliman. 2014. Impacts of marine invaders on biodiversity depend on trophic position and functional similarity. Marine Ecology Progress Series **495**:39-47.

Thomsen, P. F., and E. E. Sigsgaard. 2019. Environmental DNA metabarcoding of wild flowers reveals diverse communities of terrestrial arthropods. Ecology and Evolution **9**:1665-1679.

Tsuji, S., T. Takahara, H. Doi, N. Shibata, and H. Yamanaka. 2019. The detection of aquatic macroorganisms using environmental DNA analysis—A review of methods for collection, extraction, and detection. Environmental DNA **1**:99-108.

Tsuji, S., M. Ushio, S. Sakurai, T. Minamoto, and H. Yamanaka. 2017. Water temperature-dependent degradation of environmental DNA and its relation to bacterial abundance. PLoS One **12**:e0176608.

Turner, C. R., M. A. Barnes, C. C. Y. Xu, S. E. Jones, C. L. Jerde, and D. M. Lodge. 2014. Particle size distribution and optimal capture of aqueous macrobial eDNA. Methods in Ecology and Evolution **5**:676-684.

Turner, C. R., K. L. Uy, and R. C. Everhart. 2015. Fish environmental DNA is more concentrated in aquatic sediments than surface water. Biological Conservation **183**:93-102.

Uthicke, S., M. Lamare, and J. R. Doyle. 2018. eDNA detection of corallivorous seastar (*Acanthaster cf. solaris*) outbreaks on the Great Barrier Reef using digital droplet PCR. Coral Reefs **37**:1229-1239.

Utzeri, V. J., G. Schiavo, A. Ribani, S. Tinarelli, F. Bertolini, S. Bovo, and L. Fontanesi. 2018. Entomological signatures in honey: an environmental DNA metabarcoding approach can disclose information on plant-sucking insects in agricultural and forest landscapes. Scientific Reports **8**:9996.

van der Loos, L. M., and R. Nijland. 2020. Biases in bulk: DNA metabarcoding of marine communities and the methodology involved. Molecular Ecology.

van der Plas, F. 2019. Biodiversity and ecosystem functioning in naturally assembled communities. Biological Reviews **94**:1220-1245.

Vellend, M., M. Dornelas, L. Baeten, R. Beauséjour, C. D. Brown, P. De Frenne, S. C. Elmendorf, N. J. Gotelli, F. Moyes, I. H. Myers-Smith, A. E. Magurran, B. J. McGill, H. Shimadzu, and C. Sievers. 2017. Estimates of local biodiversity change over time stand up to scrutiny. Ecology **98**:583-590.

Vergés, A., P. D. Steinberg, M. E. Hay, A. G. Poore, A. H. Campbell, E. Ballesteros, K. L. Heck Jr, D. J. Booth, M. A. Coleman, and D. A. Feary. 2014. The tropicalization of temperate marine ecosystems: climate-mediated changes in herbivory and community phase shifts. Proceedings of the Royal Society B: Biological Sciences **281**:20140846.

Vilà, M., and P. E. Hulme. 2017. Non-native Species, Ecosystem Services, and Human Well-Being. Pages 1-14 *in* M. Vilà and P. E. Hulme, editors. Impact of Biological Invasions on Ecosystem Services. Springer International Publishing, Cham.

Wagner, D. L. 2020. Insect declines in the Anthropocene. Annual review of entomology **65**:457-480.

Wallace, A., R. 1876. The geographical distribution of animals, with a study of the relations of living and extinct faunas as elucidating the past changes of the earth's surface. Macmillan and Co, London.

Walther, G.-R., A. Roques, P. E. Hulme, M. T. Sykes, P. Pyšek, I. Kühn, M. Zobel, S. Bacher, Z. Botta-Dukát, H. Bugmann, B. Czúcz, J. Dauber, T. Hickler, V. Jarošík, M. Kenis, S. Klotz, D. Minchin, M. Moora, W. Nentwig, J. Ott, V. E. Panov, B. Reineking, C. Robinet, V. Semenchenko, W. Solarz, W. Thuiller, M. Vilà, K. Vohland, and J. Settele. 2009. Alien species in a warmer world: risks and opportunities. Trends in Ecology & Evolution **24**:686-693.

Walton, W. C., C. MacKinnon, L. F. Rodriguez, C. Proctor, and G. M. Ruiz. 2002. Effect of an invasive crab upon a marine fishery: green crab, Carcinus maenas, predation upon a venerid clam, Katelysia scalarina, in Tasmania (Australia). Journal of Experimental Marine Biology and Ecology **272**:171-189.

Waycott, M., C. M. Duarte, T. J. B. Carruthers, R. J. Orth, W. C. Dennison, S. Olyarnik, A. Calladine, J. W. Fourqurean, K. L. Heck, A. R. Hughes, G. A. Kendrick, W. J. Kenworthy, F. T. Short, and S. L. Williams. 2009. Accelerating loss of seagrasses across the globe threatens coastal ecosystems. Proceedings of the National Academy of Sciences **106**:12377.

Weigle, S. M., L. D. Smith, J. T. Carlton, and J. Pederson. 2005. Assessing the risk of introducing exotic species via the live marine species trade. Conservation Biology **19**:213-223.

Westfall, K. M., T. W. Therriault, and C. L. Abbott. 2020. A new approach to molecular biosurveillance of invasive species using DNA metabarcoding. Global Change Biology **26**:1012-1022.

Wilcox, T. M., K. S. McKelvey, M. K. Young, W. H. Lowe, and M. K. Schwartz. 2015. Environmental DNA particle size distribution from Brook Trout (*Salvelinus fontinalis*). Conservation Genetics Resources **7**:639-641.

Wilcox, T. M., K. E. Zarn, M. P. Piggott, M. K. Young, K. S. McKelvey, and M. K. Schwartz. 2018. Capture enrichment of aquatic environmental DNA: A first proof of concept. Molecular Ecology Resources **18**:1392-1401.

Williams, J. M., C. A. Duckworth, M. D. Burkitt, A. J. Watson, B. J. Campbell, and D. M. Pritchard. 2015. Epithelial cell shedding and barrier function: a matter of life and death at the small intestinal villus tip. Vet Pathol **52**:445-455.

Williams, M. A., J. O'Grady, B. Ball, J. Carlsson, E. de Eyto, P. McGinnity, E. Jennings, F. Regan, and A. Parle-McDermott. 2019. The application of CRISPR-Cas for single species identification from environmental DNA. Molecular Ecology Resources **19**:1106-1114.

Williams, M. R., R. D. Stedtfeld, C. Engle, P. Salach, U. Fakher, T. Stedtfeld, E. Dreelin, R. J. Stevenson, J. Latimore, and S. A. Hashsham. 2017. Isothermal amplification of environmental DNA (eDNA) for direct field-based monitoring and laboratory confirmation of *Dreissena sp*. PLoS One **12**:e0186462.

Williams, S. L., I. C. Davidson, J. R. Pasari, G. V. Ashton, J. T. Carlton, R. E. Crafton, R. E. Fontana, E. D. Grosholz, A. W. Miller, G. M. Ruiz, and C. J. Zabin. 2013. Managing Multiple Vectors for Marine Invasions in an Increasingly Connected World. BioScience **63**:952-966.

Williamson, M., and A. Fitter. 1996. The Varying Success of Invaders. Ecology **77**:1661-1666.

Willis, A. 2016. Extrapolating abundance curves has no predictive power for estimating microbial biodiversity. Proceedings of the National Academy of Sciences **113**:E5096.

Wilson, E. O. 1988. Editor's Foreword. Page 538 *in* E. O. Wilson and H. University, editors. Biodiversity. The National Academies Press, Washington, DC.

Wood, S. A., X. Pochon, W. Ming, U. von Ammon, C. Woods, M. Carter, M. Smith, G. Inglis, and A. Zaiko. 2018. Considerations for incorporating real-time PCR assays into routine marine biosecurity surveillance programmes: a case study targeting the Mediterranean fanworm (*Sabella spallanzanii*) and club tunicate (*Styela clava*). Genome **62**:137-146.

Worm, B., and H. K. Lotze. 2021. Chapter 21 - Marine biodiversity and climate change. Pages 445-464 *in* T. M. Letcher, editor. Climate Change (Third Edition). Elsevier.

Wotton, R. S., and B. Malmqvist. 2001. Feces in Aquatic Ecosystems: Feeding animals transform organic matter into fecal pellets, which sink or are transported horizontally by currents; these fluxes relocate organic matter in aquatic ecosystems. BioScience **51**:537-544.

Yamamoto, S., K. Minami, K. Fukaya, K. Takahashi, H. Sawada, H. Murakami, S. Tsuji, H. Hashizume, S. Kubonaga, T. Horiuchi, M. Hongo, J. Nishida, Y. Okugawa, A. Fujiwara, M. Fukuda, S. Hidaka, K. W. Suzuki, M. Miya, H.

Araki, H. Yamanaka, A. Maruyama, K. Miyashita, R. Masuda, T. Minamoto, and M. Kondoh. 2016. Environmental DNA as a 'Snapshot' of Fish Distribution: A Case Study of Japanese Jack Mackerel in Maizuru Bay, Sea of Japan. PLoS One 11:e0149786.

Yates, M. C., D. J. Fraser, and A. M. Derry. 2019. Meta-analysis supports further refinement of eDNA for monitoring aquatic species-specific abundance in nature. Environmental DNA 1:5-13.

Zulkefli, N. S., K.-H. Kim, and S.-J. Hwang. 2019. Effects of microbial activity and environmental parameters on the degradation of extracellular environmental DNA from a eutrophic lake. International Journal of Environmental Research and Public Health 16:3339.

Chapter 2 Detection of introduced and resident marine species using environmental DNA metabarcoding of sediment and water

2.1 Citation

Here I present a full typeset article as published in *Scientific reports*, the citation is provided below.

Holman LE, de Bruyn M, Creer S, Carvalho G, Robidart, J, Rius M (2019) The detection of novel and resident marine non-indigenous species using environmental DNA metabarcoding of seawater and sediment. Scientific Reports 9:11559. DOI:10.1038/s41598-019-47899-7

www.nature.com/scientificreports

SCIENTIFIC REPORTS
natureresearch

Corrected: Publisher Correction

OPEN

Detection of introduced and resident marine species using environmental DNA metabarcoding of sediment and water

Received: 2 January 2019
Accepted: 25 July 2019
Published online: 09 August 2019

Luke E. Holman[1], Mark de Bruyn[2,3], Simon Creer[2], Gary Carvalho[2], Julie Robidart[4] & Marc Rius[1,5]

Environmental DNA (eDNA) surveys are increasingly being used for biodiversity monitoring, principally because they are sensitive and can provide high resolution community composition data. Despite considerable progress in recent years, eDNA studies examining how different environmental sample types can affect species detectability remain rare. Comparisons of environmental samples are especially important for providing best practice guidance on early detection and subsequent mitigation of non-indigenous species. Here we used eDNA metabarcoding of COI (cytochrome c oxidase subunit I) and 18S (nuclear small subunit ribosomal DNA) genes to compare community composition between sediment and water samples in artificial coastal sites across the United Kingdom. We first detected markedly different communities and a consistently greater number of distinct operational taxonomic units in sediment compared to water. We then compared our eDNA datasets with previously published rapid assessment biodiversity surveys and found excellent concordance among the different survey techniques. Finally, our eDNA surveys detected many non-indigenous species, including several newly introduced species, highlighting the utility of eDNA metabarcoding for both early detection and temporal / spatial monitoring of non-indigenous species. We conclude that careful consideration on environmental sample type is needed when conducting eDNA surveys, especially for studies assessing community change.

Anthropogenic activities have widespread impacts on global biodiversity[1,2] and can negatively affect ecosystem services and function[3]. Cumulatively these actions create an urgent need to develop monitoring tools that rapidly and accurately detect community composition in ecosystems. Existing biodiversity survey techniques have been criticised for their methodological limitations (e.g. observer bias or taxonomic resolution)[4,5] and are typically standardised by a survey time limit or through reaching asymptote of a species discovery curve[6,7]. Such surveys often focus on the detection of a specific taxonomic group that are being targeted, with no ability to retrospectively separate mis-identified species in light of new species discoveries. This is of critical importance for biodiversity monitoring as an increasing number of studies are revealing the widespread presence of molecular cryptic species (i.e. morphologically similar but genetically distinct species[8]). For example, between 9,000–35,000 marine species (2.7% of the total number of known marine species) are considered molecular cryptic, and genetic studies often reveal widespread marine species containing multiple cryptic lineages[9,10]. This highlights the need to integrate morphological and genetic approaches to accurately detect community composition.

One approach that has the potential to overcome some of the above limitations is the use of DNA found in environmental samples, such as water, soil or sediment, to infer presence or absence of organisms in the ecosystem[11]. This genetic material, known as environmental DNA (eDNA), is a poly disperse mixture of tissue, cells,

[1]School of Ocean and Earth Science, National Oceanography Centre Southampton, University of Southampton, Southampton, United Kingdom. [2]Molecular Ecology and Fisheries Genetics Laboratory, School of Natural Sciences, Bangor University, Bangor, United Kingdom. [3]The University of Sydney, School of Life and Environmental Sciences, Sydney, Australia. [4]Ocean Technology and Engineering Group, National Oceanography Centre Southampton, Southampton, United Kingdom. [5]Centre for Ecological Genomics and Wildlife Conservation, University of Johannesburg, Johannesburg, South Africa. Correspondence and requests for materials should be addressed to L.E.H. (email: l.e.holman@soton.ac.uk)

subcellular fragments and extracellular DNA lost to the environment through the normal life and death of organisms[12,13]. Environmental DNA surveys have been used in targeted detection (i.e. single species) studies with qPCR assays[14–17], and in community (i.e. multi-species) studies using metabarcoding[18–20]. These surveys are highly sensitive and once the methodology is optimised are amenable to automation[21,22]. However, validity and replicability rely on appropriate experimental design and an understanding of the effects of methodological choices during sampling, sequencing library preparation and bioinformatic analysis[23,24]. Although it is well-established that eDNA surveys are highly informative and can complement other biodiversity monitoring methods[25], eDNA studies assessing how different sampling techniques affect species detectability remain rare[26].

An area where accurate monitoring tools are critical is the detection of non-indigenous species (NIS). NIS are those that have been transported through human action from their native range into a novel geographic location. Only a subset of the total number of NIS have a net negative effect[27] but these pose a severe threat to anthropogenic activities, human health and indigenous biodiversity[28–32]. Most marine NIS have spread globally via vectors such as transoceanic shipping or canals connecting large water bodies[32,33]. At smaller (tens of km) geographical scales, other vectors such as intraregional boating significantly enhance the spread of NIS[34]. In coastal areas, studies have highlighted the importance of monitoring marinas and harbours[6], as these are hotspots of NIS and together with marine infrastructure (e.g. breakwaters, artificial reefs) promote the spread of NIS[35]. In these habitats, NIS often outcompete native resident species and dominate artificial hard substrata[36,37]. Marinas and harbours have distinct ecological and physico-chemical conditions compared to the surrounding marine environment[38,39]. Consequently, specific sampling and surveying protocols are needed to study marine organisms in these environments, with eDNA surveys offering huge potential for early detection and management of NIS.

Recent work has identified a vast range of protocols for the collection and extraction of eDNA from different environmental sample types (e.g. water, sediment)[40–42]. Despite this progress, we are only just beginning to understand how the choice of environmental sample type affects species detectability[26,43]. For example, we would not expect to detect nektonic in addition to benthic organisms in an analysis of a sediment core using microscopy, but several eDNA studies have detected both of these groups in eDNA isolated from marine sediment[26,44]. Understanding which proportion of the total community is detected using eDNA isolated from different sample types is essential to place eDNA surveys in the context of existing methods, especially when studying NIS.

Here we used eDNA metabarcoding of COI (cytochrome c oxidase subunit I) and 18S (nuclear small subunit ribosomal DNA) genes to compare alpha and beta diversity between sediment and water samples collected in marine urban environments. We then compared the eDNA metabarcoding results with previously published biodiversity data to identify if NIS detection was comparable between methods. We subsequently parsed our eDNA metabarcoding dataset to identify NIS in the study region. We then outlined the strengths and weaknesses of eDNA metabarcoding for the detection of NIS and more broadly community composition. Finally, we discussed how this technique can help conservation efforts for both assessing indigenous biodiversity and mitigating the deleterious effects of NIS.

Results

Raw sequencing results and taxonomic annotation.

Sequencing produced a total of 17.8 million paired end reads, with 15.2 million sequences remaining after paired end read merging and quality filtering. The average number of sequences per sample after filtering (excluding those from control samples) was 200,185 ± 64,019 (s.d.). Negative control samples contained an average of 811 ± 3,402 (s.d.) sequences. One negative control sample contained ~15,000 sequences that mapped to an operational taxonomic unit (OTU) that had 100% identity to a sequence of a terrestrial fungi (Genbank Accession number: FJ804151.1). Excluding this OTU from the entire analysis resulted in an average of 51 ± 94 (s.d.) sequences per no-template control sample. Denoising produced 8,069 OTUs for COI and 2,433 for 18S with 6,435 and 1,679 remaining respectively after OTU curation with LULU. Taxonomic annotation identified 622 OTUs from the 18S rRNA dataset and 481 OTUs from the COI dataset. Taxonomic data from World Register of Marine Species[45] could be retrieved for 200 of the annotated COI OTUs and 190 of the 18S OTUs.

OTU richness and community structure.

The effects of preservation techniques for water eDNA samples differed between the target amplicons. The 18S rRNA amplicon produced significantly more OTUs (Wilcoxon signed-rank test, $p < 0.05$) in samples preserved by freezing compared to Longmire's preservation method, while no significant differences (Wilcoxon signed-rank test, $p = 0.55$) were observed between preservation treatments for the COI amplicon (see Supplementary Information 2 for details). As a conservative approach all subsequent analyses used sample data from the frozen samples. The minimum number of reads per sample was 137,624 and 117,915 for the COI and 18S datasets, respectively, and so samples were rarefied to these numbers of reads. A consistently greater number of OTUs were detected in the sediment samples compared to the water samples across all sites and both markers as shown in Fig. 1b,c. In all cases, unique OTUs were detected in both water and sediment samples, but the mean proportion of unique OTUs across 18S and COI detected in water was lower (49.2%) than in sediment (73.8%). A two-way ANOVA testing the effects of sample type, site and their interaction on the number of OTUs indicated a significant effect of the site-sample type interaction ($p < 0.001$) for both 18S and COI (see Supplementary Information 3 for full model output). Ordination plots based on the Bray-Curtis dissimilarities (Fig. 1d,e) showed that OTUs found in sediment and water eDNA differed in community structure as much as among sites. Additionally, the PERMANOVA model indicated significant differences ($p < 0.001$) among sites and eDNA sample types in both the 18S and COI datasets (see Supplementary Information 4 for full model output). Accordingly, eDNA sample type in the PERMANOVA model explained 23.2% and 32.5% of the variation in the 18S and COI data respectively, while site explained 34.2% and 30.5% in the 18S and COI data. Species detections binned at Phylum level showed variable detection sample type within Phylum (Fig. 2). However, an exact binomial goodness of fit test showed non-random detection proportions in Nematoda and Platyhelminthes ($p < 0.001$

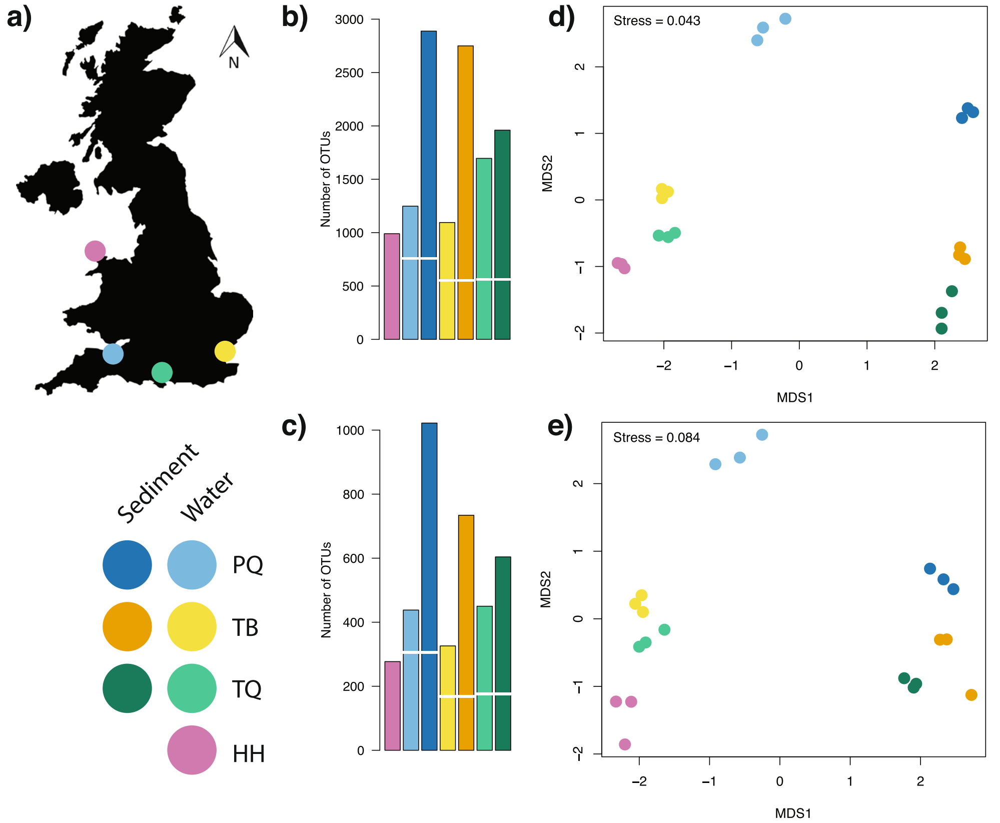


Figure 1. (**a**) Map of the United Kingdom indicating the geographic position of the sampled sites, a legend indicates the four sites (PQ, TB, TQ and HH) and colours for water and sediment eDNA samples for each site. Barplots detailing number of OTUs (operational taxonomic units) detected across sampling sites and eDNA sample type for (**b**) COI and (**c**) 18S rRNA metabarcoding of UK marinas, the break in bars indicates the number of shared OTUs between sediment and water eDNA samples. Non-metric multidimensional scaling ordination plots based on Bray-Curtis dissimilarities of: (**d**) COI and (**e**) 18S rRNA metabarcoding of marina sediment and water eDNA samples.

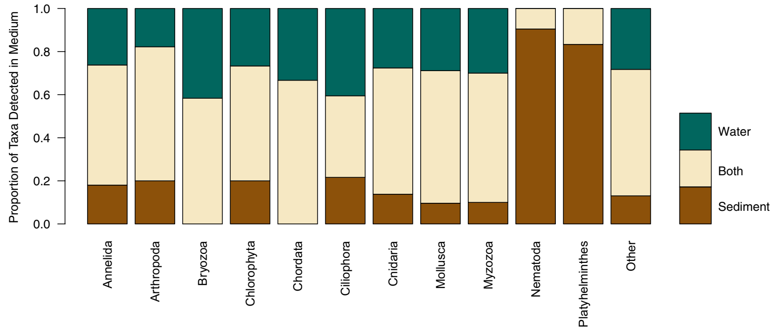


Figure 2. Horizontal stacked barchart detailing proportion of operational taxonomic units detected in eDNA from sediment, water or both sediment and water across the 14 Phyla for pooled 18S rRNA and COI metabarcoding data from the sampled marinas.

and $p = 0.038$ respectively, see Supplementary Information 5 for full details), with species detections mostly in sediment in both cases.

Detection of non-indigenous species. As the 18S region lacks the appropriate resolution for taxonomic assignments at species level[46,47] only the taxonomic assignments from the COI were considered for the identification of NIS. In total 18 NIS to the study region and 24 species documented as NIS in other regions were detected across the four sites (see Supplementary Table 2 for full list). Out of the detected NIS, eight were present in the

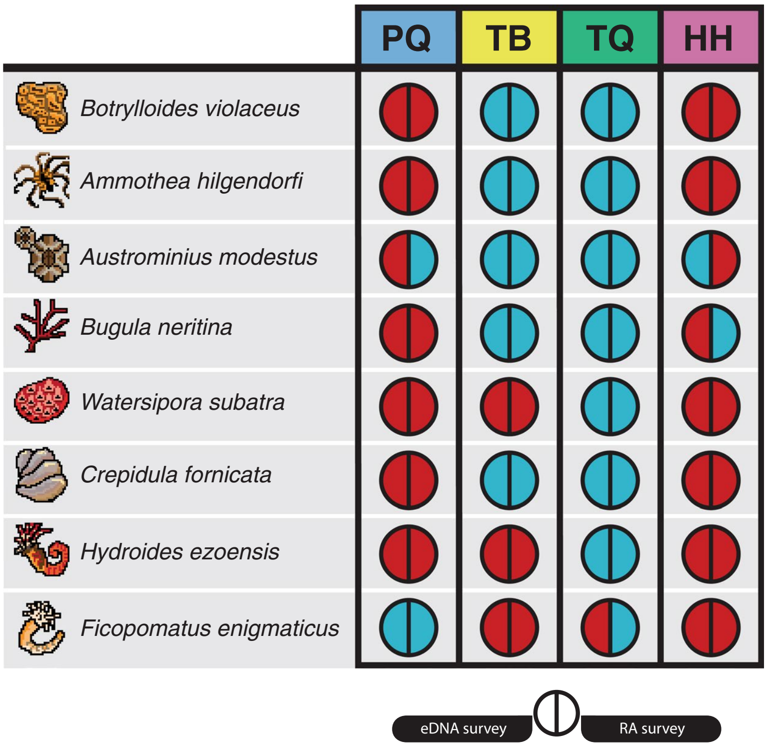


Figure 3. Incidence diagram for eight non-indigenous species across the four sampling sites (PQ,TB,TQ and HH). For each species-location the left semi-circle indicates the detection during our eDNA metabarcoding surveys of 18S rRNA and COI fragments, and the right semi-circle indicates the detection from rapid assessment (RA) surveys. Blue indicates a positive detection for that species-location and red indicates no detection.

list of 21 NIS previously detected in rapid assessment (RA) surveys at the sampling sites[48–50]. As shown in Fig. 3, the results of the eDNA surveys closely matched those of the RA surveys. Four detections differed from the RA surveys, a single eDNA detection not seen in RA and three RA detections not seen in eDNA surveys (Fig. 3). Remapping of raw reads from sites with incongruent detections to respective COI regions (Genbank Accessions: *Austrominius modestus* KY607884; *Bugula neritina* KY235450; *Ficopomatus enigmatus* KX840011) found hits for the bryozoan *B. neritina* only (five reads from a single replicate). These reads were lost during data filtering and so did not feature in the final dataset. Three species detections in site TQ represented recent introductions: the detection of *Arcuatula senhousia* (Asian date mussel), the nemertean *Cephalothrix simula* and the oligochaete *Paranais frici*. Targeted visual surveys on tidal mudflats within two kilometres of Marina TQ confirmed the presence of live *A. senhousia* individuals. Furthermore, we generated COI sequences from tissue samples of these individuals (Genbank Accession: MH924820 and MH924821) and these provided full length, high identity matches to both known *A. senhousia* DNA sequences and our eDNA derived OTU sequence (see Supplementary Information 6 for details of DNA barcoding).

Discussion

We demonstrated that the type of environmental sample in eDNA metabarcoding studies affects the measured community composition, indicating that the most comprehensive assessment of biodiversity in a given community comes from the collection of multiple environmental sample types. In addition, we found concordance between our eDNA metabarcoding data and previous biodiversity surveys, demonstrating complementarity of different biodiversity assessment methods. Furthermore, we detected recently introduced NIS, providing support for eDNA metabarcoding as an effective tool for early detection of NIS. This is key as early detection of NIS greatly increases the likelihood of successful control and eventual eradication of NIS. Overall, we demonstrate that type of environmental sample can affect the detection of both whole community composition and particular species of concern.

Our study showed that taxonomic assignments at the level of Phylum did not predict if a species was detected in water, sediment or both environmental sample types (except in Nematodes and Platyhelminthes, whose members are predominantly benthic inhabitants). However, all sampled sites showed higher OTU richness in sediment compared to water. The magnitude of this difference was not fixed across sites, with a significant interaction term in our two-way ANOVA (Supplementary Information 3) indicating that the detected OTU richness differences between sediment and water vary spatially. The majority of research using eDNA to detect aquatic macrofauna has focused on the collection of water samples, while sediment samples have received comparatively less attention (see Fig. S1 from Koziol, *et al.*[26]). This is surprising considering that sediment samples typically contain three orders of magnitude more eDNA than water[51]. Despite our observations that sediment provided a greater number of OTU richness than water samples, we do not advocate for a particular sample type, as this decision should be driven by the target organisms for a given study. For example, a researcher hoping to use eDNA metabarcoding to measure Nematode diversity, based on our results, should sample marine sediment. Regarding NIS, both water and sediment served as excellent sample types for NIS detection. Consequently, our results suggest that no specific sample type offers a better detection of NIS, likely because NIS are not found in a single phylogenetic clade' for clarity. We argue that at a lower taxonomic level, the species-specific ecology of eDNA (*sensu* Barnes and Turner[52]) may lead to convergent eDNA occupancy in different environmental sample types. Further work is needed to clarify how eDNA partitions into adjacent environmental samples across the tree of life. A key unknown is the underlying explanation for eDNA metabarcoding data from sediment samples generating more OTUs in comparison to water. One hypothesis is that eDNA from sediment includes extracellular 'free' DNA that is not retained by the filters used to process water for eDNA samples. Studies focussing on eDNA surveys have found little evidence identifying what proportion of total eDNA is extracellular DNA. However, using qPCR Turner and colleagues[12] identified that eDNA particles with a size of less than 0.2µm, well below the size of intra-organellular DNA, are less than 10% of the total eDNA pool for a teleost fish species. If this pattern is observed in other metazoans then extracellular eDNA may have little effect on the differences of OTU richness detected here. An alternative hypothesis is that due to eDNA settlement and persistence dynamics in sediment, it contains a greater diversity of eDNA fragments (both extra and intracellular).

Current eDNA metabarcoding research has identified large variation in the detected marine biodiversity across small spatial scales (hundreds of metres) in both sediment[53] and water[54,55]. Additionally, fractionation of environmental samples (i.e. sorting samples by particle size class) can produce significant differences in the metabarcoding results between fractions[56,57] indicating significant variation within sites. We found similar patterns, with PERMANOVA modelling showing approximately equivalent variation in OTU dissimilarity between site and environmental sample type. Future research should explore how different sample types and eDNA extraction methods affect the detection of marine species, especially as eDNA metabarcoding moves from an experimental technique to a routine monitoring tool[58,59].

A key gap in our understanding is the rate at which eDNA degrades in sediment and how this affects our observations. In lake sediments, eDNA can be preserved for thousands of years[60,61], with eDNA being preserved along with deposited sediments so each core represents a timeline through which past biological communities can be examined[62]. Here we chose to process only the uppermost section of the sampled cores, with the aim of profiling contemporary species composition. Studies are needed to advance our understanding of how eDNA deposits and degrades in marine sediments in order to temporally contextualise sediment samples.

We found that eDNA metabarcoding accurately detected many NIS, as seen in previous studies[18–20]. By comparing our eDNA data to those collected using existing methods we found close congruence in NIS incidence. The false-negative eDNA detection of *B. neritina* was found to be a result of setting specific bioinformatic parameters, showing that choices made during sequence processing can have a significant effect on the detectability of species in eDNA samples. Indeed, this has previously been shown in metabarcoding of bulk tissue samples[63] and work is urgently needed to determine the effects of bioinformatic parameters, variable primer binding sites and the choice of reference databases on the detection of NIS from eDNA samples. The remaining incongruent detections may be a result of community turnover among the survey dates or phenological changes affecting species distributions. Indeed, marine coastal communities have been shown to shift in community composition across seasons and reproductive cycles[64,65]. Therefore, our data suggest that in order to enhance existing monitoring programmes, replicated eDNA metabarcoding surveys over time should be performed.

In our study we identified several recently introduced NIS in the United Kingdom and confirmed the eDNA detection with targeted local surveys for one NIS. The case of *A. senhousia* is particularly relevant as it is spreading globally[66] and has the potential to dramatically alter benthic biodiversity when invasive[67,68]. This species produces a cocoon of byssus thread that at high densities (>1,500 individuals/m^2) interlinks between individuals to form a continuous byssal mat which displaces local eelgrass and native bivalves[69,70]. Recent field surveys along the south coast of the United Kingdom have independently confirmed the presence of both *A. senhousia*[71] and *C. simula*[72]. These results confirm the accuracy of eDNA surveys presented here and highlight the benefits of implementing molecular technologies for routine monitoring programmes.

As the cost of sequencing continues to decrease and methods improve across the metabarcoding workflow[73] natural resource managers and researchers will have access to much greater resolution data at a fraction of the cost and time of current monitoring surveys. However, NIS can be missed in surveys based solely on eDNA (e.g. Wood, *et al.*[74]) and eDNA studies can detect rare species that are often missed using other methods[75]. Detection of NIS could be further facilitated through autonomous sampling and eDNA surveys[21] to provide live species incidence data in introduction hotspots, such as ports or marinas. Additionally combining these techniques with eDNA biobanking[76] could provide an eDNA reference database for specific geographical regions of high biosecurity risk, providing an invaluable resource for both biodiversity managers and researchers to examine the process of biological invasion through time. Taken together, our study shows eDNA metabarcoding to be an effective tool for the detection and identification of both resident and recently introduced species from different environmental samples.

Methods

Study sites. Four marinas were selected from around the United Kingdom (Fig. 1a) to represent variation in modelled invasion potential[77], presence of NIS[78] and benthic habitat type[79]. All chosen marinas have been surveyed previously, so there is a good understanding of the species found in these sites[48–50]. Marina access was contingent on anonymity and so marina names and exact locations are not provided, with Fig. 1a showing approximate locations only. Marina *TQ* is an open marina subject to tides and varying salinity, marina *PQ* is a loch marina open during high tide, and marinas *TB* and *HH* are permanently open to the North Sea and Celtic Sea respectively.

Environmental DNA sampling. Surveys were conducted during May 2017 (see Supplementary Table 1 for site details) and 24 sampling points were randomly selected within each site. At each sampling point 50 ml of water was collected from 10 cm below the surface using a sterile 60 ml Luer lock syringe and filtered through a 0.22 μm polyethersulfone Sterivex filter (Merck Millipore, Massachusetts USA). After collecting seawater from eight sampling points (400 ml total volume) the filter was changed, resulting in a total of three filters per site. Pooling of water samples was performed to provide three filter replicates per site that represented the heterogeneity of eDNA in the marina. In order to test the effect of different sample preservation methods, water samples were collected in duplicate in each sampling point. One set of three filters had ~1.5 ml sterile Longmire's solution (100 mM Tris,10 mM EDTA, 10 mM NaCl, 0.5% SDS) applied in the inlet valve[80]. The second set of three filters was kept on ice for no longer than eight hours before being frozen at −20 °C. In addition to the water samples, a subtidal sediment sample was collected at the first water sampling point and then after every three water samples, accounting for a total of nine sediment samples per site. A UWITEC Corer (UWITEC, Mondsee, Austria) was used to collect a sediment core of 600 mm high and 60 mm diameter. A sterile disposable spatula was used to collect a subsample of 10–20 g of sediment from the top 2 cm of the core, avoiding sediment collection from the sides of the core. The subsamples were stored in sterile plastic bags and kept on ice for no longer than eight hours before being frozen at −80 °C. Due to a malfunction of the corer, no sediment sample was collected in Site HH. Disposable gloves were changed after collection of each sample. All reused equipment was soaked in 10% bleach and rinsed in DNase-free sterile water between sites.

eDNA extraction. DNA extractions were performed in a PCR-free clean room, separate from main laboratory facilities. No high copy templates, cultures or amplicons were permitted in this clean laboratory. DNA extractions from water samples followed the $\mathrm{SX}_{\mathrm{CAPSULE}}$ method in Spens, *et al.*[41]. Briefly, preservative solution was removed from the outlet and filters were dried at room temperature for two hours. 720 μl Qiagen buffer ATL (Qiagen, Hilden, Germany) and 80 μl Proteinase K (20 mg/ml) was added to the filter and all samples were digested overnight at 56 °C. After digestion, samples were processed using the Qiagen DNeasy Blood and Tissue Kit as per manufacturer instructions, with a final elution of 200 μl PCR grade water.

Sediment extractions were conducted using the Qiagen DNeasy Powermax Soil Kit following the manufacturer's protocol. The nine samples collected in each site were randomly mixed to form three pooled samples; 10 g of pooled sample was processed for the extraction. A total of ten samples were processed, three from each site with a single extraction control.

Inhibition testing. To ensure extracted DNA was free of PCR inhibitors, a Primer Design Real-Time PCR Internal Control Kit (PrimerDesign, Southampton, United Kingdom) was used. qPCR reactions were performed for each sample following the manufacturer's protocol with 12.5 μl reaction volumes containing 2 μl of extracted eDNA sample. A positive detection of inhibition due to co-purified compounds from DNA extraction protocols would produce an increase in cycle threshold number (>1.0) in comparison to no template controls. All samples were successfully processed and no samples showed indication of PCR inhibition.

Primer selection and library preparation. Two sets of primers were chosen for metabarcoding the environmental samples: a 313 bp section of the standard DNA barcoding region of the cytochrome c oxidase subunit I gene using primers described in Leray, *et al.*[81]; and a variable length target of the hypervariable V4 region of the nuclear small subunit ribosomal DNA using primers from Zhan, *et al.*[82]. These two primer sets allow for broad characterisation of marine metazoan diversity. Sequencing libraries were prepared using a 2-step PCR approach as detailed in Bista, *et al.*[83]. Briefly, this method first amplifies the target region in PCR 1 annealing universal adapters, and then sample specific indices and sequencing primers are annealed in PCR 2. In contrast to Bista, *et al.*[83] we used unique dual-matched indexes for PCR 2 to avoid index crosstalk associated with combinatorial indexing[84]. PCR 1 was prepared in a PCR–free room separate from main laboratory facilities. PCR 1 reactions were conducted in 20 μl volumes containing 10 μl Amplitaq Gold 360 2X Mastermix (Applied Biosystems, California, USA), 0.8 μl (5 nmol ml^{-1}) of each forward and reverse primer and 2 μl of undiluted environmental DNA extract. The reaction conditions for PCR were an initial denaturation step at 95 °C for 10 minutes followed by 20 cycles of 95 °C for 0:30, variable annealing temp (46 °C for COI and 50 °C for 18S) for 0:30, and extension at 72 °C for 1:00. A final extension at 72 °C was performed for 10 minutes. The PCR product was cleaned using AMPure XP beads (Beckman Coulter, California, USA) at a 0.8 beads:sample ratio following manufacturer's instructions. PCR 2 reactions were conducted in 20 μl volumes containing 10 μl Amplitaq GOLD 360 2X Mastermix, 0.5 μl (10 nmol ml^{-1}) of both forward and reverse primers and 5 μl of undiluted cleaned PCR1 product. PCR conditions were an initial denaturation step at 95 °C for 10 minutes followed by 15 cycles of 95 °C for 0:30, annealing at 55 °C for 0:30, and extension at 72 °C for 1:00. A final extension at 72 °C was performed for 10 minutes. PCR 2 products were cleaned using AMpure XP beads as above and normalised according to their fluorescence using the Qubit HS Assay Kit (Thermofisher Scientific, Massachusetts, USA). These normalised samples were pooled at an equimolar concentration and then quantified as per manufacturer's instructions using the NEBNext Library Quant qPCR kit (New England Biolabs, Massachusetts, USA).

Blank filters, DNA extraction kits and positive controls were collected, extracted and sequenced identically to non-control samples (detailed in Supplementary Information 1). Negative controls cannot be meaningfully normalized and thus they were added to the pooled libraries without dilution. The final library was sequenced using an Illumina MiSeq instrument (Illumina, San Diego, USA) with a V3 2 × 300 bp kit.

Bioinformatic analyses. Samples were demultiplexed using the Illumina MiSeq control software (v.2.6.2.1). The demultiplexed data was analysed using a custom pipeline written in the R programming language[85] (hosted at https://github.com/leholman/metabarTOAD). The steps are as follows. Forward and reverse paired end reads were merged using the -fastq_mergepairs option of USEARCH v.10.0.240[86] with maximum difference of 15, percent identity of 80% and quality filter set at maximum expected errors of 1. Both the forward and reverse primer sequences were matched using Cutadapt v.1.16[87] and only sequences containing both primer regions were retained. Sequences were discarded if they were outside of a defined length boundary (303–323 bp for COI, 375–450 bp for 18S) using Cutadapt. Sequences were then pooled, singletons were discarded and sequences were quality filtered with a maximum expected error of 1 using the -fastq_filter option of VSEARCH v.2.4.3[88]. Sequences were then denoised and chimeras filtered using the unoise3 algorithm implemented in USEARCH. The resultant operational taxonomic units (OTUs) were curated using the LULU package v.0.1.0[89]. An OTU by sample table was produced by mapping the merged and trimmed reads against the curated OTUs using USEARCH, with the raw query read assigned to the OTU with the best match (highest bit score) within 97% identity. The OTU by sample table was filtered in R (v.3.5.0) as follows. To minimise the chance of spurious OTUs being included in the final dataset any record with less than 3 raw reads were changed to zero and any OTU that did not appear in more than one sample was removed from the analysis. OTUs found in negative controls were removed from the analysis.

Taxonomic assignment. Assigning correct taxonomy to an unknown set of DNA sequences can be challenging as reference databases are incomplete, contain errors and the taxonomy of some marine groups is uncertain. With such limitations in mind, we assigned taxonomy using a BLAST v.2.6.0+ search[90] returning the single best hit (largest bit score) from databases within 97% of the query using a custom R script to parse the raw blast results. In the case of multiple sequences attaining equal bit scores for a given OTU an assignment was only made if all reference sequences belonged to the same species. The MIDORI database (UNIQUE_20180221)[91] was used for the COI data and the SILVA database (SSU r132, subset to contain only Eukaryotes)[92] was used for the 18S rRNA data. The match taxa tool from the World Register of Marine Species[45] was used to filter the data to include only marine species and check the taxonomic classification. The World Register of Introduced Marine Species[93] contains a range of peer-reviewed and technical reports on the global introduced status of a large number of species, we used the online match taxa tool to determine the non-indigenous status of annotations that could be assigned taxonomy from the World Register of Marine Species.

Statistical analyses. All statistical analyses were conducted in R v.3.5.0. The Vegan R package [94] was used to rarefy samples to the minimum sample read depth for each amplicon. The number of OTUs per site/condition was calculated as the number of OTUs with a non-zero number of normalized reads after summing the reads across all three site level replicates. To test if there was a significant difference between the number of OTUs generated by sediment and water eDNA, individual non-summed replicate sample data was used to build a two-way ANOVA model with the formula *number_of_OTUs~sedimentorwater*site* implemented in R using the function *aov*. Non-metric multidimensional scaling ordination plots were generated from Bray-Curtis dissimilarity values derived using vegan. A Permutation Analysis of Variance (PERMANOVA)[95] was performed using the Bray-Curtis dissimilarity following the model *dissimilarity_matrix~sedimentorwater*site* implemented in R using the function *adonis* from the vegan package. OTUs with taxonomic assignment were separated into those found in sediment, water or both media and the OTUs were then collapsed at the Phylum level to explore taxonomic patterns of detection in water or sediment. Phyla with less than eight OTUs were combined and represented under category named "other". To test for non-random counts of species detection between water and sediment within taxa an exact binomial test was performed between counts of species detected in water and sediment. The number of species detected in both water and sediment were halved and the value added to the counts for each sample type with non-integer values conservatively rounded down to the nearest whole number. A correction for multiple comparisons[96] was applied across the p values from the exact binomial tests generated by the R function *binom.test*. Records from rapid assessment surveys previously conducted for non-native invertebrates at the sample sites[48–50] were compared with the detected species from metabarcoding data.

Data Availability

Raw Illumina sequencing data is available from the European Nucleotide Archive under under study accession number PRJEB33619. Associated metadata, R scripts and intermediate files are available online via Zenodo with the following https://doi.org/10.5281/zenodo.1453958.

References

1. Sala, E. & Knowlton, N. Global marine biodiversity trends. *Annual Review of Environment and Resources* **31**, 93–122, https://doi.org/10.1146/annurev.energy.31.020105.100235 (2006).
2. Butchart, S. H. M. *et al.* Global biodiversity: indicators of recent declines. *Science* **328**, 1164–1168, https://doi.org/10.1126/science.1187512 (2010).
3. Worm, B. *et al.* Impacts of biodiversity loss on ocean ecosystem services. *Science* **314**, 787–790, https://doi.org/10.1126/science.1132294 (2006).
4. Oliver, I. & Beattie, A. J. A possible method for the rapid assessment of biodiversity. *Conservation Biology* **7**, 562–568, https://doi.org/10.1046/j.1523-1739.1993.07030562.x (1993).
5. Fitzpatrick, M. C., Preisser, E. L., Ellison, A. M. & Elkinton, J. S. Observer bias and the detection of low-density populations. *Ecological Applications* **19**, 1673–1679 (2009).

6. Ashton, G., Books, K., Shucksmith, R. & Cook, E. Rapid assessment of the distribution of marine non-native species in marinas in Scotland. *Aquatic Invasions* **1**, 209–213 (2006).
7. Bishop, J. D., Wood, C. A., Leveque, L., Yunnie, A. L. & Viard, F. Repeated rapid assessment surveys reveal contrasting trends in occupancy of marinas by non-indigenous species on opposite sides of the western English Channel. *Marine Pollution Bulletin* **95**, 699–706, https://doi.org/10.1016/j.marpolbul.2014.11.043 (2015).
8. Appeltans, W. *et al*. The magnitude of global marine species diversity. *Current Biology* **22**, 2189–2202, https://doi.org/10.1016/j.cub.2012.09.036 (2012).
9. Pérez-Portela, R., Arranz, V., Rius, M. & Turon, X. Cryptic speciation or global spread? The case of a cosmopolitan marine invertebrate with limited dispersal capabilities. *Scientific Reports* **3**, 3197, https://doi.org/10.1038/srep03197 (2013).
10. Rius, M. & Teske, P. R. Cryptic diversity in coastal Australasia: a morphological and mitonuclear genetic analysis of habitat-forming sibling species. *Zool J Linn Soc-Lond* **168**, 597–611, https://doi.org/10.1111/zoj.12036 (2013).
11. Thomsen, P. F. & Willerslev, E. Environmental DNA - an emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation* **183**, 4–18, https://doi.org/10.1016/j.biocon.2014.11.019 (2015).
12. Turner, C. R. *et al*. Particle size distribution and optimal capture of aqueous macrobial eDNA. *Methods in Ecology and Evolution* **5**, 676–684, https://doi.org/10.1111/2041-210x.12206 (2014).
13. Sassoubre, L. M., Yamahara, K. M., Gardner, L. D., Block, B. A. & Boehm, A. B. Quantification of environmental DNA (eDNA) shedding and decay rates for three marine fish. *Environmental Science & Technology* **50**, 10456–10464, https://doi.org/10.1021/acs.est.6b03114 (2016).
14. Dougherty, M. M. *et al*. Environmental DNA (eDNA) detects the invasive rusty crayfish *Orconectes rusticus* at low abundances. *Journal of Applied Ecology* **53**, 722–732, https://doi.org/10.1111/1365-2664.12621 (2016).
15. Simpson, T. J. S., Dias, P. J., Snow, M., Muñoz, J. & Berry, T. Real-time PCR detection of *Didemnum perlucidum* (Monniot, 1983) and *Didemnum vexillum* (Kott, 2002) in an applied routine marine biosecurity context. *Molecular Ecology Resources* **17**, 443–453, https://doi.org/10.1111/1755-0998.12581 (2017).
16. Wood, S. A., Zaiko, A., Richter, I., Inglis, G. J. & Pochon, X. Development of a real-time polymerase chain reaction assay for the detection of the invasive Mediterranean fanworm, *Sabella spallanzanii*, in environmental samples. *Environmental Science and Pollution Research* **24**, 17373–17382, https://doi.org/10.1007/s11356-017-9357-y (2017).
17. Kim, P., Kim, D., Yoon, T. J. & Shin, S. Early detection of marine invasive species, *Bugula neritina* (Bryozoa: Cheilostomatida), using species-specific primers and environmental DNA analysis in Korea. *Marine Environmental Research* **139**, 1–10, https://doi.org/10.1016/j.marenvres.2018.04.015 (2018).
18. Borrell, Y. J., Miralles, L., Do Huu, H., Mohammed-Geba, K. & Garcia-Vazquez, E. DNA in a bottle-rapid metabarcoding survey for early alerts of invasive species in ports. *PLos One* **12**, e0183347, https://doi.org/10.1371/journal.pone.0183347 (2017).
19. Lacoursiere-Roussel, A. *et al*. eDNA metabarcoding as a new surveillance approach for coastal Arctic biodiversity. *Ecology and Evolution* **8**, 7763–7777, https://doi.org/10.1002/ece3.4213 (2018).
20. Grey, E. K. *et al*. Effects of sampling effort on biodiversity patterns estimated from environmental DNA metabarcoding surveys. *Scientific Reports* **8**, 8843, https://doi.org/10.1038/s41598-018-27048-2 (2018).
21. McQuillan, J. S. & Robidart, J. C. Molecular-biological sensing in aquatic environments: recent developments and emerging capabilities. *Current Opinion in Biotechnology* **45**, 43–50, https://doi.org/10.1016/j.copbio.2016.11.022 (2017).
22. Yamahara, K. M. *et al*. *In situ* Autonomous Acquisition and Preservation of Marine Environmental DNA Using an Autonomous Underwater Vehicle. *Front Mar Sci* **6**, https://doi.org/10.3389/fmars.2019.00373 (2019).
23. Goldberg, C. S. *et al*. Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution* **7**, 1299–1307, https://doi.org/10.1111/2041-210x.12595 (2016).
24. Alberdi, A., Aizpurua, O., Gilbert, M. T. P. & Bohmann, K. Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution* **9**, 134–147, https://doi.org/10.1111/2041-210x.12849 (2018).
25. Deiner, K. *et al*. Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology* **26**, 5872–5895, https://doi.org/10.1111/mec.14350 (2017).
26. Koziol, A. *et al*. Environmental DNA metabarcoding studies are critically affected by substrate selection. *Molecular Ecology Resources* **19**, 366–376, https://doi.org/10.1111/1755-0998.12971 (2019).
27. Anton, A. *et al*. Global ecological impacts of marine exotic species. *Nature Ecology & Evolution* **3**, 787–800, https://doi.org/10.1038/s41559-019-0851-0 (2019).
28. Bax, N., Williamson, A., Aguero, M., Gonzalez, E. & Geeves, W. Marine invasive alien species: a threat to global biodiversity. *Marine Policy* **27**, 313–323, https://doi.org/10.1016/S0308-597x(03)00041-1 (2003).
29. Lovell, S., Stone, S. & Fernandez, L. The Economic Impacts of Aquatic Invasive Species: A Review of the Literature. *Agricultural and Resource Economics Review* **35**, 195–208, https://doi.org/10.1017/S1068280500010157 (2006).
30. Ricciardi, A., Hoopes, M. F., Marchetti, M. P. & Lockwood, J. L. Progress toward understanding the ecological impacts of nonnative species. *Ecological Monographs* **83**, 263–282, https://doi.org/10.1890/13-0183.1 (2013).
31. Mazza, G., Tricarico, E., Genovesi, P. & Gherardi, F. Biological invaders are threats to human health: an overview. *Ethology Ecology &. Evolution* **26**, 112–129, https://doi.org/10.1080/03949370.2013.863225 (2014).
32. Molnar, J. L., Gamboa, R. L., Revenga, C. & Spalding, M. D. 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 (2008).
33. Nunes, A. L., Katsanevakis, S., Zenetos, A. & Cardoso, A. C. Gateways to alien invasions in the European seas. *Aquatic Invasions* **9**, 133–144, https://doi.org/10.3391/ai.2014.9.2.02 (2014).
34. Murray, C. C., Pakhomov, E. A. & Therriault, T. W. Recreational boating: a large unregulated vector transporting marine invasive species. *Diversity and Distributions* **17**, 1161–1172, https://doi.org/10.1111/j.1472-4642.2011.00798.x (2011).
35. Airoldi, L., Turon, X., Perkol-Finkel, S. & Rius, M. Corridors for aliens but not for natives: effects of marine urban sprawl at a regional scale. *Diversity and Distributions* **21**, 755–768, https://doi.org/10.1111/ddi.12301 (2015).
36. Glasby, T. M., Connell, S. D., Holloway, M. G. & Hewitt, C. L. Nonindigenous biota on artificial structures: could habitat creation facilitate biological invasions? *Marine Biology* **151**, 887–895, https://doi.org/10.1007/s00227-006-0552-5 (2006).
37. Dafforn, K. A., Johnston, E. L. & Glasby, T. M. Shallow moving structures promote marine invader dominance. *Biofouling* **25**, 277–287, https://doi.org/10.1080/08927010802710618 (2009).
38. Rivero, N. K., Dafforn, K. A., Coleman, M. A. & Johnston, E. L. Environmental and ecological changes associated with a marina. *Biofouling* **29**, 803–815, https://doi.org/10.1080/08927014.2013.805751 (2013).
39. Foster, V., Giesler, R. J., Wilson, A. M. W., Nall, C. R. & Cook, E. J. Identifying the physical features of marina infrastructure associated with the presence of non-native species in the UK. *Marine Biology* **163**, 163–173, https://doi.org/10.1007/s00227-016-2941-8 (2016).
40. Deiner, K. *et al*. Optimising the detection of marine taxonomic richness using environmental DNA metabarcoding: the effects of filter material, pore size and extraction method. *Metabarcoding and Metagenomics* **2**, e28963, https://doi.org/10.3897/mbmg.2.28963 (2018).
41. Spens, J. *et al*. Comparison of capture and storage methods for aqueous macrobial eDNA using an optimized extraction protocol: advantage of enclosed filter. *Methods in Ecology and Evolution* **8**, 635–645, https://doi.org/10.1111/2041-210x.12683 (2017).
42. Sellers, G. S., Di Muri, C., Gómez, A. & Hänfling, B. Mu-DNA: a modular universal DNA extraction method adaptable for a wide range of sample types. *Metabarcoding and Metagenomics* **2**, e24556, https://doi.org/10.17504/protocols.io.qn9dvh6) (2018).

43. Hermans, S. M., Buckley, H. L. & Lear, G. Optimal extraction methods for the simultaneous analysis of DNA from diverse organisms and sample types. *Molecular Ecology Resources* **18**, 557–569, https://doi.org/10.1111/1755-0998.12762 (2018).
44. Shaw, J. L. A. *et al*. Comparison of environmental DNA metabarcoding and conventional fish survey methods in a river system. *Biological Conservation* **197**, 131–138, https://doi.org/10.1016/j.biocon.2016.03.010 (2016).
45. WoRMS Editorial Board. *World Register of Marine Species*, https://doi.org/10.14284/170, http://www.marinespecies.org (2019).
46. Leray, M. & Knowlton, N. Censusing marine eukaryotic diversity in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences* **371**, 20150331, https://doi.org/10.1098/rstb.2015.0331 (2016).
47. Mohrbeck, I., Raupach, M. J., Martinez Arbizu, P., Knebelsberger, T. & Laakmann, S. High-throughput sequencing - the key to rapid biodiversity assessment of marine metazoa? *PLoS One* **10**, e0140342, https://doi.org/10.1371/journal.pone.0140342 (2015).
48. Wood, C. A., Bishop, J. D. D. & Yunnie, A. L. E. RAS 2014: non-native species rapid assessment surveys in English marinas. 34pp (2015).
49. Wood, C. A., Bishop, J. D. D. & Yunnie, A. L. E. *Comprehensive Reassessment of NNS in Welsh Marinas*. 42pp (2015).
50. Wood, C. A., Bishop, J. D. D., Rennocks, L. & Crundwell, R. RAS 2015: non-native species rapid assessment surveys in English marinas (E Anglia & W coast). 34pp (2016).
51. Torti, A., Lever, M. A. & Jorgensen, B. B. Origin, dynamics, and implications of extracellular DNA pools in marine sediments. *Marine Genomics* **24**(3), 185–196, https://doi.org/10.1016/j.margen.2015.08.007 (2015).
52. Barnes, M. A. & Turner, C. R. The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics* **17**, 1–17, https://doi.org/10.1007/s10592-015-0775-4 (2016).
53. Nascimento, F. J. A., Lallias, D., Bik, H. M. & Creer, S. Sample size effects on the assessment of eukaryotic diversity and community structure in aquatic sediments using high-throughput sequencing. *Scientific reports* **8**, 11737, https://doi.org/10.1038/s41598-018-30179-1 (2018).
54. O'Donnell, J. L. *et al*. Spatial distribution of environmental DNA in a nearshore marine habitat. *PeerJ* **5**, e3044, https://doi.org/10.7717/peerj.3044 (2017).
55. Jeunen, G. J. *et al*. Environmental DNA (eDNA) metabarcoding reveals strong discrimination among diverse marine habitats connected by water movement. *Molecular Ecology Resources* **19**, 426–438, https://doi.org/10.1111/1755-0998.12982 (2019).
56. Wangensteen, O. S., Palacin, C., Guardiola, M. & Turon, X. 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 (2018).
57. Wangensteen, O. S. & Turon, X. In *Marine Animal Forests: The Ecology of Benthic* Biodiversity Hotspots (eds S. Rossi, L. Bramanti, A. Gori, & C. Orejas) Ch. Metabarcoding techniques for assessing biodiversity of marine animal forests, 445–473 (Springer, 2017).
58. Pawlowski, J. *et al*. The future of biotic indices in the ecogenomic era: Integrating (e)DNA metabarcoding in biological assessment of aquatic ecosystems. *Science of the Total Environment* **637–638**, 1295–1310, https://doi.org/10.1016/j.scitotenv.2018.05.002 (2018).
59. Aylagas, E., Borja, A., Muxika, I. & Rodríguez-Ezpeleta, N. 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 (2018).
60. Ficetola, G. F. *et al*. DNA from lake sediments reveals long-term ecosystem changes after a biological invasion. *Science Advances* **4**, eaar4292, https://doi.org/10.1126/sciadv.aar4292 (2018).
61. Pedersen, M. W. *et al*. Postglacial viability and colonization in North America's ice-free corridor. *Nature* **537**, 45, https://doi.org/10.1038/nature19085 (2016).
62. Balint, M. *et al*. Environmental DNA time series in ecology. *Trends in Ecology & Evolution* **33**, 945–957, https://doi.org/10.1016/j.tree.2018.09.003 (2018).
63. Scott, R. *et al*. Optimization and performance testing of a sequence processing pipeline applied to detection of nonindigenous species. *Evolutionary Applications* **11**, 891–905, https://doi.org/10.1111/eva.12604 (2018).
64. Stachowicz, J. J. & Byrnes, J. E. Species diversity, invasion success, and ecosystem functioning: disentangling the influence of resource competition, facilitation, and extrinsic factors. *Marine Ecology Progress Series* **311**, 251–262, https://doi.org/10.3354/meps311251 (2006).
65. Sutherland, J. P. & Karlson, R. H. Development and stability of the fouling community at Beaufort, North Carolina. *Ecological Monographs* **47**, 425–446 (1977).
66. Bachelet, G. *et al*. A round-the-world tour almost completed: first records of the invasive mussel *Musculista senhousia* in the north-east. *Atlantic (southern Bay of Biscay). Marine Biodiversity Records* **2**, 2002–2005, https://doi.org/10.1017/s1755267209001080 (2009).
67. Crooks, J. A. Assessing invader roles within changing ecosystems: historical and experimental perspectives on an exotic mussel in an urbanized lagoon. *Biological Invasions* **3**, 23–36 (2001).
68. Mistri, M. The non-indigenous mussel *Musculista senhousia* in an Adriatic lagoon: Effects on benthic community over a ten year period. *Journal of the Marine Biological Association of the United Kingdom* **83**, 1277–1278, https://doi.org/10.1017/S0025315403008658 (2003).
69. Kushner, R. B. & Hovel, K. A. Effects of native predators and eelgrass habitat structure on the introduced Asian mussel *Musculista senhousia* (Benson in Cantor) in southern California. *Journal of Experimental Marine Biology and Ecology* **332**, 166–177, https://doi.org/10.1016/j.jembe.2005.11.011 (2006).
70. Crooks, J. A. & Khim, H. S. Architectural vs. biological effects of a habitat-altering, exotic mussel, *Musculista senhousia*. *Journal of Experimental Marine Biology and Ecology* **240**, 53–75, https://doi.org/10.1016/S0022-0981(99)00041-6 (1999).
71. Barfield, P., Holmes, A., Watson, G. & Rowe, G. First evidence of *Arcuatula senhousia* (Benson, 1842), the asian date mussel in UK waters. *Journal of Conchology* **43**, 217–222 (2018).
72. Turner, A. D. *et al*. New invasive nemertean species (*Cephalothrix simula*) in England with high levels of tetrodotoxin and a microbiome linked to toxin metabolism. *Mar Drugs* **16**, 452, https://doi.org/10.3390/md16110452 (2018).
73. Elbrecht, V. & Steinke, D. Scaling up DNA metabarcoding for freshwater macrozoobenthos monitoring. *Freshwater Biol* **64**, 380–387, https://doi.org/10.1111/fwb.13220 (2019).
74. Wood, S. A. *et al*. Considerations for incorporating real-time PCR assays into routine marine biosecurity surveillance programmes: a case study targeting the Mediterranean fanworm (*Sabella spallanzanii*) and club tunicate (*Styela clava*). *Genome* **62**, 137–146, https://doi.org/10.1139/gen-2018-0021 (2019).
75. Blackman, R. C. *et al*. Detection of a new non-native freshwater species by DNA metabarcoding of environmental samples – first record of *Gammarus fossarum* in the UK. *Aquatic Invasions* **12**, 177–189, https://doi.org/10.3391/ai.2017.12.2.06 (2017).
76. Jarman, S. N., Berry, O. & Bunce, M. The value of environmental DNA biobanking for long-term biomonitoring. *Nature Ecology and Evolution* **2**, 1192–1193, https://doi.org/10.1038/s41559-018-0614-3 (2018).
77. Pearce, F., Peeler, E. & Stebbing, P. Modelling the risk of the introduction and spread of non-indigenous species in the UK and Ireland. *Project Report for E5405W. CEFAS* (2012).
78. Bishop, J. D. D., Wood, C. A., Yunnie, A. L. E. & Griffiths, C. A. Unheralded arrivals: non-native sessile invertebrates in marinas on the English coast. *Aquatic Invasions* **10**, 249–264, https://doi.org/10.3391/ai.2015.10.3.01 (2015).
79. Calewaert, J. B., Weaver, P., Gunn, V., Gorringe, P. & Novellino, A. In Quantitative Monitoring of the Underwater Environment: Results of the International Marine Science and Technology Event MOQESM '14 in Brest, France (eds Zerr, B. *et al*.) 31–46 (Springer International Publishing, 2016).

80. Renshaw, M. A., Olds, B. P., Jerde, C. L., McVeigh, M. M. & Lodge, D. M. The room temperature preservation of filtered environmental DNA samples and assimilation into a phenol-chloroform-isoamyl alcohol DNA extraction. *Molecular Ecology Resources* **15**, 168–176, https://doi.org/10.1111/1755-0998.12281 (2015).
81. Leray, M. *et al.* 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 (2013).
82. Zhan, A. *et al.* High sensitivity of 454 pyrosequencing for detection of rare species in aquatic communities. *Methods in Ecology and Evolution* **4**, 558–565, https://doi.org/10.1111/2041-210x.12037 (2013).
83. Bista, I. *et al.* Annual time-series analysis of aqueous eDNA reveals ecologically relevant dynamics of lake ecosystem biodiversity. *Nature. Communications* **8**, 14087, https://doi.org/10.1038/ncomms14087 (2017).
84. MacConaill, L. E. *et al.* Unique, dual-indexed sequencing adapters with UMIs effectively eliminate index cross-talk and significantly improve sensitivity of massively parallel sequencing. *BMC Genomics* **19**, 30, https://doi.org/10.1186/s12864-017-4428-5 (2018).
85. R_Core_Team. R: a language and environment for statistical computing. *ISBN 3-900051-07-0* (2018).
86. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods* **10**, 996–998, https://doi.org/10.1038/nmeth.2604 (2013).
87. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal* **17**, 10–12 (2011).
88. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahe, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**, e2584, https://doi.org/10.7717/peerj.2584 (2016).
89. Frøslev, T. G. *et al.* Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nature. Communications* **8**, 1188, https://doi.org/10.1038/s41467-017-01312-x (2017).
90. Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421, https://doi.org/10.1186/1471-2105-10-421 (2009).
91. Machida, R. J., Leray, M., Ho, S. L. & Knowlton, N. Metazoan mitochondrial gene sequence reference datasets for taxonomic assignment of environmental samples. *Scientific Data* **4**, 170027, https://doi.org/10.1038/sdata.2017.27 (2017).
92. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* **41**, D590–596, https://doi.org/10.1093/nar/gks1219 (2013).
93. Ahyong, S. *et al. World Register of Introduced Marine Species (WRiMS)*, https://doi.org/10.14284/347, www.marinespecies.org/introduced (2019).
94. Oksanen, J. *et al.* Vegan: community ecology package. *R package* **1**, 17 (2011).
95. Anderson, M. J. In *Wiley Stats Ref: Statistics Reference Online* (eds Balakrishnan, N. *et al.*) 1–15 (John Wiley & Sons, Ltd, 2014).
96. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J R Stat Soc B* **57**, 289–300 (1995).

Acknowledgements

We are grateful to John Bishop and Chris Wood from the Marine Biological Association of the United Kingdom for sharing information on marinas and their excellent NIS survey data. We thank the staff of the Environmental Sequencing Facility from the National Oceanography Centre Southampton for advice and assistance during library preparation. We thank Dr Ivan Haigh for assistance in accessing remote sensing data. We acknowledge the Department of Geography and Environment from the University of Southampton for access to coring equipment and laboratory space. LH was supported by the Natural Environmental Research Council (grant number NE/L002531/1).

Author Contributions

L.E.H. and M.R. designed the experiment, L.E.H. collected samples, generated and analysed the data, prepared all figures and wrote the first draft of the paper. L.E.H., M.B., S.C., G.C., J.R. and M.R. substantially contributed to further manuscript drafts.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-47899-7.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019

2.3 Supplementary Information

2.3.1 Supplementary Information 1

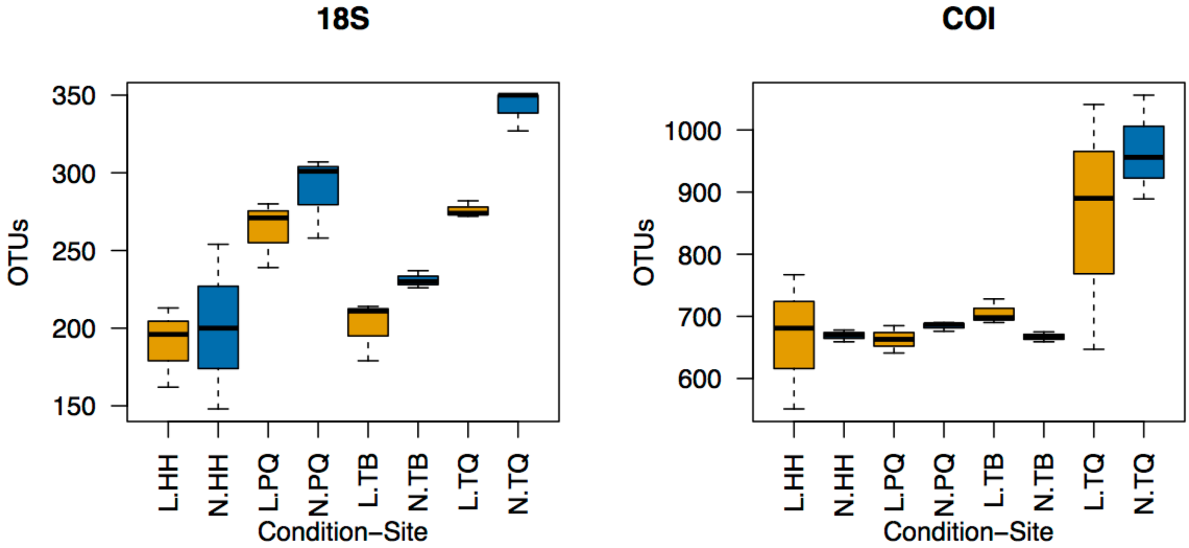
Description of control samples used in experimental procedures.

The following control samples were used. Two sealed filter controls were taken during all field sampling and subject to identical treatment as samples taken during each sampling trip, one contained Longmire's solution (100 mM Tris, 100 mM EDTA, 10 mM NaCl, 0.5 % SDS), the second was kept cool during sampling. A filter was opened during DNA extraction to act as an equipment blank for all filters. Ultrapure water (400ml) was left open in the post-PCR lab for 3 weeks then filtered, this control checked for contaminating amplicons in aerosols in the lab. A blank extraction was used for the DNeasy kit following the same protocol but with no template added. PCR no-template controls were run during PCR1 and PCR2. A positive control sample of 200ml was filtered from a marine tropical aquarium in the reception of the National Oceanography Centre, Southampton, United Kingdom (latitude 50.891380, longitude -1.3939209). The tank contained a variety of hard and soft corals, molluscs and tropical fish. A 400ml sample of seawater was filtered adjacent to the National Oceanography Centre Southampton and 5µl of 1:100 diluted extracted genomic DNA from a species not currently known to United Kingdom waters; *Microcosmus squamiger* (Class Ascidiacea, Phylum Chordata) was added to the filter before DNA extraction to act as an inhibition control. This control was used to test if target DNA could be detected after DNA extraction in the presence of inhibitors found in marinas and harbours.

2.3.2 Supplementary Information 2

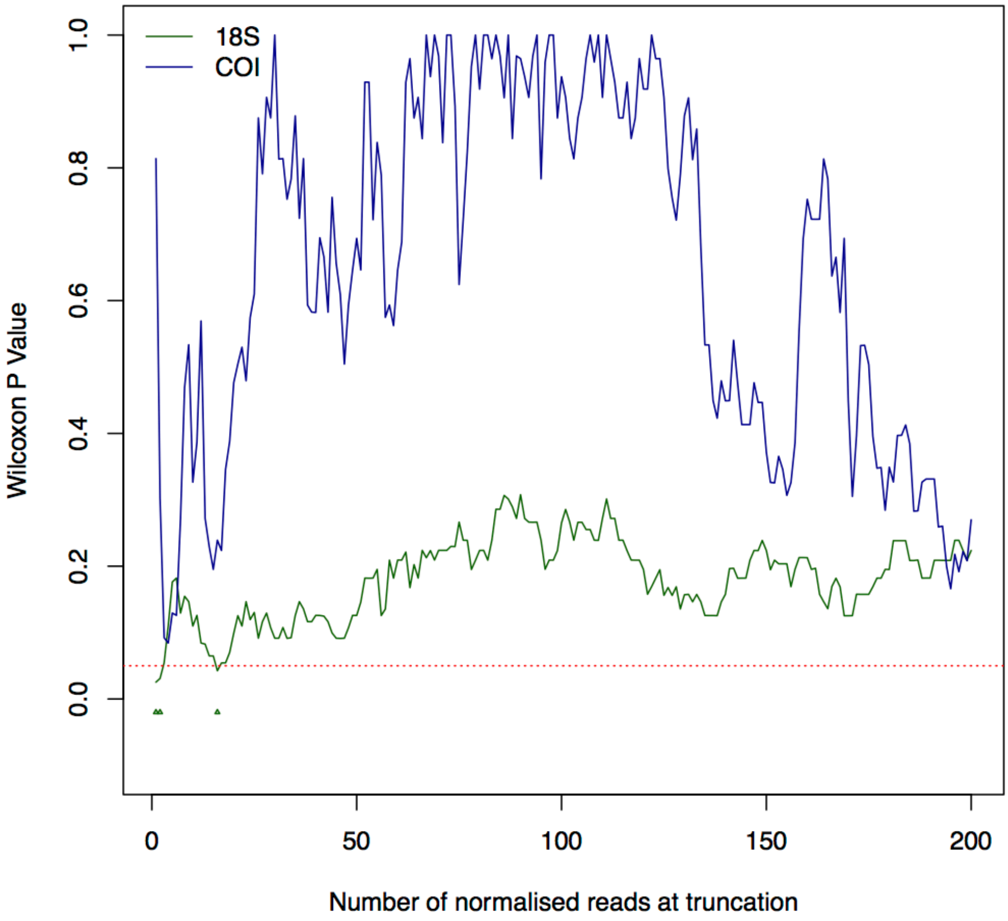
Details of analyses used to compare different methods for environmental DNA sample preservation.

Raw data was rarefied by the lowest number of reads per sample, this was 117,915 in the 18S dataset and 52,740 in the COI dataset. Detection of an OTU was positive if a non-zero number of normalised reads mapped to an OTU. As shown in Figure S1 below the average number of OTUs detected was always higher in the frozen sample compared to the Longmire's preserved sample in the 18S data, there was no consistent difference between conditions in the COI dataset. The data was tested for statistical significance using a Wilcoxon signed-rank test. This test showed more OTUs are detected in water samples preserved by freezing in comparison to Longmire's solution using a 18S amplicon (V=10, p=0.025). No significant difference was found in OTU detection between sample preservation methods using a COI amplicon (V=38, p=0.969).



Supplementary Information 2 Figure 1 *Boxplots detailing number of OTUs generated using eDNA metabarcoding of seawater in marina sites with samples preserved with either Longmire's solution (Yellow) or by Freezing (Blue).*

In order to examine if the effect between preservation methods is driven by low abundance OTUs the analysis was rerun with OTU detection being positive at an increasing threshold of normalised reads from 1 to 200. The results of the Wilcoxon signed-rank test for a significant difference between Longmire's and frozen OTU detection are shown in Figure S2 below. The results indicate that the significant effect is driven by low abundance OTUs in the 18S dataset. After applying the Bonferroni correction for multiple comparison no significant difference remained between the conditions.



Supplementary Information 2 Figure 2 *Line chart detailing Wilcoxon signed-rank test P value for a difference between the number of OTUs detected across pairs of samples preserved using Longmire' solution or by freezing as a function of number of normalised reads at truncation. The coloured solid lines indicate the values for a 18S and COI amplicon. The dashed red line marks a value of 0.05 and the coloured points indicate significance at the P<0.05 level.*

Taken together, these results indicate that there is a difference in the number of OTUs generated in an eDNA metabarcoding experiment between Longmire's and temperature preservation when using a 18S amplicon. This difference is driven by low abundance OTUs which may represent unfiltered false-positive detections or rare sequences.

2.3.3 Supplementary Information 3

*Table containing model output for linear model with formula number_of_OTUs~sedimentorwater*site for both 18S and COI metabarcoding of UK marinas*

COI

	Df	Sum Sq	Mean Sq	F Value	Pr(>F)
cond	1	4285616	4285616	171.254	1.83E-08
site	2	229337	114668	4.582	0.0332
cond:site	2	1322855	661428	26.431	4.01E-05
Residuals	12	300299	25025		

18S

	Df	Sum Sq	Mean Sq	F Value	Pr(>F)
cond	1	393089	393089	98.29	3.93E-07
site	2	89435	44717	11.18	0.001813
Cond:site	2	118605	59303	14.83	0.000571
Residuals	12	47989	3999		

2.3.4 Supplementary Information 4:

*Table containing model output for PERMANOVA model with formula dissimilarity_matrix~sedimentorwater*site for both 18S and COI metabarcoding of UK Marinas*

COI

	Df	SumsOfS	MeanSqs	F.Model	R2	Pr(>F)
treatment	1	2.0771	2.07708	33.96	3.25E-01	0.001
sites	2	1.9477	0.97385	15.922	0.30508	0.001
treatment:sites	2	1.6254	0.81271	13.288	0.2546	0.001
Residuals	12	0.734	0.06116	0.11496		
Total	17	6.3842	1			

18S

	Df	SumsOfSqs	MeanSqs	F.Model	R2	Pr(>F)
treatment	1	1.5626	1.5626	16.6369	0.23213	0.001
sites	2	2.2988	1.14941	12.2377	0.3415	0.001
Treatment:sites	2	1.7431	0.87155	9.2793	0.25894	0.001
Residuals	12	1.1271	0.09392	0.16743		
Total	17	6.7316	1			

2.3.5 Supplementary Information 5

Protocol for DNA isolation and sequencing of Asian date mussels (Arcuatula senhousia).

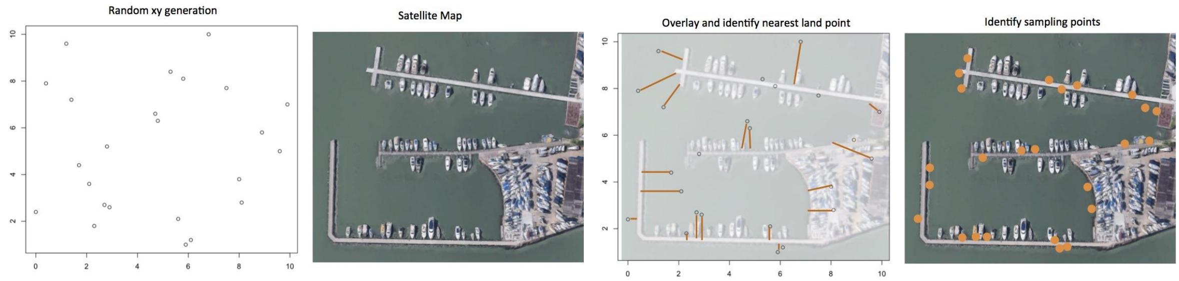
Tissue was sampled and processed in duplicate from the mantle from two *A. senhousia* individuals and immediately subject to DNA extraction using the Qiagen DNeasy Blood and Tissue Kit as manufacturer's instructions. Each sample was amplified using a set of primers targeting the COI gene (Folmer et al., 1994) conducted in 20µl volumes containing 10µl Amplitaq GOLD 360 2X Mastermix, 0.8µl (5 nmol ml^{-1}) of each forward and reverse primer and 2µl of undiluted DNA extract. The reaction conditions for PCR were an initial denaturation step at 95°C for 10 minutes followed by 35 cycles of 95°C for 30 seconds, 50°C for 30 seconds, and 72°C for 1 minute, a final extension at 72°C was performed for 10 minutes. Samples were cleaned using ExoSAP-IT Express (Applied Biosystems, California, USA) as manufacturer's instructions. Successful PCR amplicons were Sanger sequenced with both primers at Eurofins Genomics (Ebersberg, Germany). The resulting sequences were quality trimmed and aligned. Only one primer provided good quality sequencing results (LCO1490). Both concatenated sequences provided full length (>90% query cover) excellent match quality (>97% identity) BLAST hits to multiple (>20) sequences corresponding to *A. senhousia* on the NCBI nt database corresponding with many independent studies (Genbank Accessions:AB498016.1, AY570034.1, HG005372.1, HQ891034.1).

2.3.6 Supplementary Information 6

Overview of method used to select sampling location within sites.

Sampling location within sites was selected as shown in the below diagram. Random points in 2 dimensional space were selected using the R function sample(). These points were the overlaid on a map of the selected marina. The closest pontoon to each randomly selected point was then selected as a sampling point and this final map used as a guide in the field.

Supplementary Information 6 Figure 1 *Process of identifying locations within a marina for eDNA sample collection. (From left to right) Random XY co-ordinates are generated; a satellite map for the site is downloaded; random XY points are overlayed on the satellite image; the nearest possible sampling point to each random XY point is identified and used for sample collection.*


Random xy generation
Satellite Map
Overlay and identify nearest land point
Identify sampling points
0
2
4
6
8
10

Chapter 3 Animals, protists and bacteria share marine biogeographic patterns

3.1 Citation

Here I present an author's accepted manuscript, this is a post-peer-review version of an article published in *Nature Ecology & Evolution*. The final authenticated version is available online at: http://dx.doi.org/10.1038/s41559-021-01439-7, the full citation is below.

Holman LE, de Bruyn M, Creer S, Carvalho G, Robidart, J, Rius M (2021) Animals, protists and bacteria share marine biogeographic patterns. Nature Ecology & Evolution. 5 (6), 738–746. DOI:10.1038/s41598-019-47899-7

3.2 Abstract

Over millennia, ecological and evolutionary mechanisms have shaped macroecological patterns across the tree of life. Research describing these patterns at both regional and global scales has traditionally focussed on the study of metazoan species. Consequently, there is a limited understanding of cross-phyla biogeographic structuring, and an escalating need to understand the macroecology of both microscopic and macroscopic organisms. Here we used environmental DNA (eDNA) metabarcoding to explore the biodiversity of marine metazoans, protists and bacteria along an extensive and highly heterogeneous coastline. Our results showed remarkably consistent biogeographic structure across the kingdoms of life despite billions of years of evolution. Analyses investigating the drivers of these patterns for each taxonomic kingdom found that environmental conditions, such as temperature, and to a lesser extent, anthropogenic stressors such as fishing pressure and pollution, explained some of the observed variation. Additionally, metazoans displayed biographic patterns that suggested regional biotic homogenisation. Against the backdrop of global pervasive anthropogenic environmental change, our work highlights the importance of considering multiple domains of life to understand the maintenance and drivers of marine biodiversity patterns across broad taxonomic, ecological and geographical scales.

3.3 Introduction

Researchers have long recognised the importance of grouping biota into distinct, geographically separated regions. Delineating these biogeographic areas is important to understand the factors shaping the range limits of species (Spalding et al. 2007), to designate key geographic areas for biodiversity conservation (Awad et al. 2002) and to predict biological responses to environmental change (Sunday et al. 2012, Pecl et al. 2017). One of the first efforts to define geographic regions of terrestrial biota were Alfred Russel Wallace's so-called 'Zoological Regions' (Wallace 1876), which included six major regions (hereafter realms) that are still recognised today (Holt et al. 2013). The drivers responsible for these geographic classifications are predominantly environmental conditions or physical barriers. Biogeographic studies have shown that deep divergence in the geographic arrangement of terrestrial biota arose as a result of plate tectonics, while shallow divergence has been most frequently attributed to climatic conditions (Ficetola et al. 2017). In aquatic ecosystems, the relative importance of biogeographic drivers is less understood, although both climatic (e.g. temperature) (Kocsis et al. 2018) and tectonic forces (Zaffosa et al. 2017) have been identified as key determinants of marine biogeographic patterns. Recent studies have partitioned the oceans into distinct ecoregions (i.e. a geographically defined area, smaller than a realm, that contains characteristic species assemblages) (Spalding et al. 2007, Costello et al. 2017), but the description of marine ecoregions has mostly considered conspicuous or well-described species. Similarly, most marine biogeographic research has focussed on readily identifiable eukaryotic species, principally metazoans (Beck et al. 2012), although considerable progress has been made in understanding global patterns of marine microbes (Sunagawa et al. 2015). In line with recent studies demonstrating strong cross-phyla interdependence (Shade et al. 2018), there is an increasing need to include prokaryotic species in our assessment of biogeographic patterns. The language of macroecology and microbial ecology is similar, both examining the incidence of species across different spatial scales, but these fields have long progressed independently. As a result, relatively few studies have explored biogeographic patterns simultaneously for both microscopic and macroscopic life, with examples of consistent and inconsistent patterns across different taxa (Shade et al. 2018, Richter et al. 2019, Djurhuus et al. 2020). Work is thus needed to explore the consistency of biogeographic breaks across different kingdoms of life.

Human driven habitat destruction, pollution and the introduction of non-native species are key drivers of recent global biodiversity change (Pecl et al. 2017) and therefore have the potential to alter geographic patterns of biota across multiple spatial scales. Cumulatively, anthropogenic stressors not only threaten vulnerable native species but also whole-community structure and function (Naeem et al. 2012, Tilman et al. 2014, Pecl et al. 2017). The magnitude and direction of human impacts are complex, with evidence for both gains and losses in local species richness across biomes (Dornelas et al. 2014, Blowes et al. 2019, Finderup Nielsen et al. 2019). However, a consistent global pattern is emerging, with a recent and rapid increase in species turnover (Blowes et al. 2019) and an associated increase in community similarity (β diversity) between two or more geographically separated sites (Dornelas et al. 2014). Incidences of increased community similarity are known as biotic homogenisation (Olden and Rooney 2006) and are driven by human activities that promote extinctions of native species and introductions of non-native species. In light of growing evidence that taxonomic, phylogenetic and functional diversity are strongly correlated (Stuart-Smith et al. 2013), the homogenisation of biological communities has the potential to negatively affect ecosystem function. It is widely accepted that even uncommon species within an ecological community can contribute significantly to ecosystem function (Mouillot et al. 2013), demonstrating the importance of studying inconspicuous species to preserve ecosystem health. Studies have shown evidence for biotic homogenisation around the globe, with examples from plants (Finderup Nielsen et al. 2019), vertebrates (Bernardo-Madrid et al. 2019) and invertebrates (Capinha et al. 2015) demonstrating alteration of terrestrial biogeographic patterns. However, many studies are of limited taxonomic scope, focussing on highly conspicuous species for which reliable data can be easily produced (Capinha et al. 2015, Bernardo-Madrid et al. 2019). Thus, most work overlooks inconspicuous species (e.g. microbes and microscopic eukaryotes), which show vastly different reproductive, demographic and dispersal patterns compared to metazoans (Shade et al. 2018), but are known to be key actors shaping the assembly of ecological communities and ultimately underpin ecosystem functioning (Azam and Malfatti 2007). Taken together, a more comprehensive characterisation of ecological communities is clearly needed when testing the role of anthropogenic activities on biogeographic patterns.

The advent of high-throughput sequencing has revolutionised our understanding of microbial life, with studies examining global patterns of prokaryotic life now increasingly common (Sunagawa et al. 2015). Moreover, the recent and rapid development of methods to infer the incidence of larger organisms using genetic material isolated from environmental samples (known as environmental DNA or eDNA) has provided an unparalleled ability to identify species across the entire tree of life (Deiner et al. 2017, Richter et al. 2019, Djurhuus et al. 2020). Together, these methods can rapidly generate standardised biodiversity data for entire communities at unprecedented resolution, thereby minimising regional and taxonomic biases. In addition, these datasets can be analysed without complete taxonomic assignment and DNA samples can be repurposed to test novel hypotheses. A common technique is to amplify DNA barcodes from eDNA and use high-throughput sequencing to produce high-resolution biodiversity data. This method (eDNA metabarcoding) has been shown to reliably detect organisms across many different ecosystems (Deiner et al. 2017), but has infrequently been applied to understand spatial patterns of biodiversity across different kingdoms of life (Sunagawa et al. 2015, Richter et al. 2019).

A unique geographic setting for testing biogeographic hypotheses is the South African coastline, where two large water masses (the Atlantic and Indian Oceans) meet, and a wide variety of abiotic and biotic conditions are found in a single region. This coastline has three well-defined coastal ecoregions bounded by the cold western boundary Benguela Current and the warm oligotrophic eastern boundary Agulhas Current. These ecoregions have been established on the basis of studies over several decades involving a number of conspicuous metazoan taxa (Emanuel et al. 1992, Awad et al. 2002). Additionally, there is evidence for human exploitation of marine resources in the region spanning thousands of years (Griffiths et al. 2004, Griffiths et al. 2010) and some areas of the coastline have been subject to heavy maritime activity for centuries (Kaluza et al. 2010). Other human activities also prevail such as the establishment of aquaculture facilities or the construction of harbours and breakwaters (Griffiths et al. 2004, Griffiths et al. 2010). Thus, the South African coastline is an ideal study system to explore the mechanisms shaping biogeographic patterns.

Here we compared the biogeography of multiple marine kingdoms of life along the diverse coastline of South Africa. We first investigated the consistency of biogeographic boundaries across metazoans, protists and bacteria using eDNA metabarcoding. We then tested to what extent these patterns could be explained by anthropogenic and natural environmental factors. We finally evaluated if there was evidence for homogenisation of ecological communities along a coastline that has been affected by human activities for centuries.

3.4 Results

3.4.1 DNA sequencing

A total of 66.25 million sequences were produced across the three sequencing runs targeting sections of the standard DNA barcoding region of cytochrome c oxidase subunit I (hereafter COI), the V4 region of the eukaryotic nuclear small subunit ribosomal DNA (hereafter 18S) and the V3-V4 hypervariable region of prokaryotic small subunit ribosomal DNA (hereafter 16S). The number of unfiltered raw reads per experimental sample ranged from 61,958 to 859,580, with an average per sample across all three markers of 347,536 ± 109,665 (*s.d.*) (see Supplementary Table 1 for further details). Negative control samples exhibited very low levels of cross-contamination (Supplementary Note 1).

3.4.2 Taxonomic assignments & alpha diversity

After performing taxonomic assignment of sequences to metazoans, protists and bacteria, taxonomically grouped datasets with largest number of observations within each marker were used in subsequent analyses (COI for metazoans, 18S for protists & 16S for bacteria). Analyses for the remaining subsets are shown in Supplementary Note 2 and were consistent with the results presented hereafter. After taxonomic assignment to phyla 1,054, 1,433 and 2,826 ASVs (amplicon sequence variants) were retained for metazoans, protist and bacteria datasets respectively. Across all taxonomically grouped datasets the majority of detected ASVs came from a small number (<5) of phyla or supergroups (Figure 1b). This pattern was consistent across sampling sites and no major changes in the identity of ASVs at phyla or supergroup level across the study region were observed (Figure 1b).

Across all markers the greatest mean ASV richness was found along the southern coast (Fig. 1). However, a one-way analysis of variance (ANOVA) showed a significant difference ($F_{2,15}=7.18$, $p=0.007$) between coastlines only in the bacterial dataset with no difference found in both the metazoan ($F_{2,15}=1.941$, $p=0.178$) and the protist datasets ($F_{2,15}=1.416$, $p=0.273$). A post-hoc Tukey test of the bacterial data (Supplementary Table 2) showed that the east and west coasts had significantly fewer ASVs compared to the south coast, but that they were not significantly different to one another in overall ASV richness.

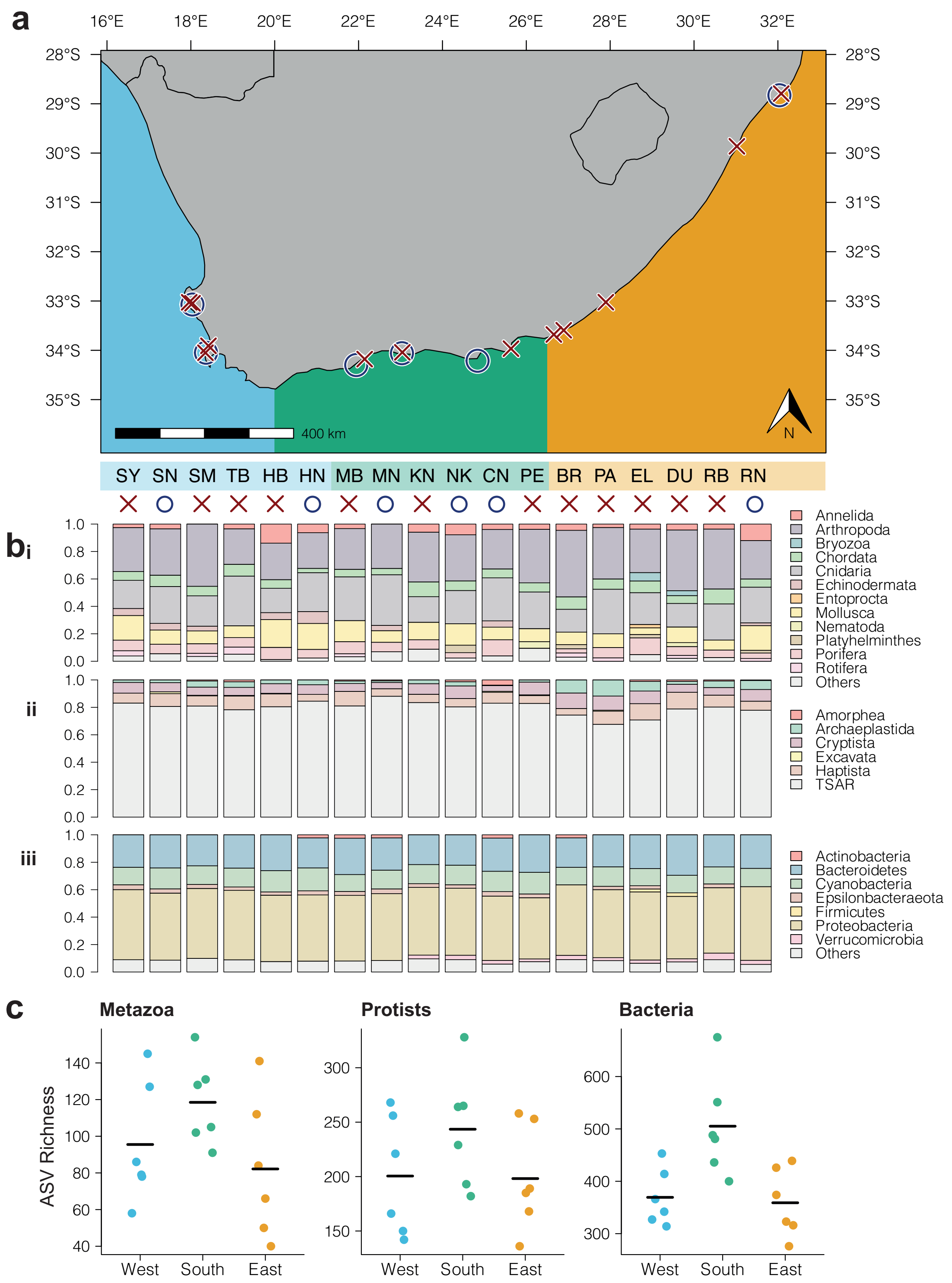


Figure 1. **a** Map of South Africa indicating the sampling sites and the site types (red crosses are artificial sites and blue circles natural sites), the east, south and west coastal regions are denoted by orange, green and blue respectively. Site codes as in Supplementary Table 9. Landmasses were plotted using the *map* function from the *maps* package in R (v3.6.1) (R_Core_Team 2019); **b** Proportion of ASVs per phyla across each site for **i** metazoans, **ii** protists and **iii** bacteria. Each bar represents a site indicated by the site code as in Supplementary Table 9; **c** Amplicon sequence variant (ASV) richness per site separated by coast (point colour matches section a) and taxonomic group, black line indicates mean ASV richness.

3.4.3 Beta diversity

Across all three taxonomic groups, non-metric multidimensional ordinations showed clustering of sites consistent with coastal ecoregions previously described in conspicuous metazoan species (Fig. 2). Furthermore, permutational multivariate analysis of variance (PERMANOVA) models showed a significant ($p<0.001$) effect of coastline in all cases (see Supplementary Table 3 for model output), with pairwise significant differences ($p<0.01$) between all pairs of coastlines in all taxa (Supplementary Table 4). There was evidence for heterogeneity of multivariate dispersion in the bacterial dataset (ANOVA on *betadispr* $F_{2,15}=4.09$, $p=0.038$, Supplementary Table 5). A Tukey test revealed a significant ($p=0.031$) pairwise difference between the east and west coast only, in line with the observations of Fig. 2., indicating that for bacteria, sites on the west coast were more variable in community composition than the more homogenous communities found at sites on the east coast. When the datasets were split by phyla, they demonstrated significant differences between ecoregions across all tested phyla ($p<0.05$ in all cases, see Supplementary Note 2 for full model output and visualisation). A power analysis (Supplementary Note 3) indicated that the number of ASVs allocated to each phylum was sufficient to detect a significant difference given the study design.

Corrected Mantel tests indicated that in the metazoan and bacterial datasets, sea surface temperature (SST) and human impact (measured by an index covering multiple anthropogenic stressors) were significantly correlated with the observed ASV dissimilarities after geographic distance between sites was accounted for (SST $p<0.05$ in all cases; human impact $p<0.01$ in all cases; full model outputs shown in Supplementary Table 6). In the protist dataset chlorophyll *a* concentration and the human impact index remained significant ($p<0.05$ in both cases). In contrast, across all datasets sea surface salinity (SSS) showed no correlation ($p>0.05$ in all cases) with observed ASV dissimilarities. These results indicated that both geographic and environmental distance have some effect on the observed community structure and also confirmed the appropriate variables to retain for analysis in each dataset. In all cases, partial Mantel tests gave similar R statistics and p values (see Supplementary Table 6).

A distanced-based redundancy analysis (dbRDA) showed a significant effect of both environmental variables ($p<0.001$ for SST and chlorophyll *a* in all cases) and human impact ($p<0.05$ in all cases) on the site similarity in metazoans, protists and bacteria (full model outputs are presented in Supplementary Table 7). Variance partitioning of the dbRDA models showed that human impact had a relatively smaller contribution to the observed dissimilarities compared to the chlorophyll *a* concentration or SSS (as shown in Fig. 2). Across taxonomic groups there was negligible overlap in the variance explained by human impact and other variables. Generalised additive models with a 2D smoothed function showed significant terms ($p>0.001$, see individual full model outputs shown in Supplementary Table 8), indicating how each variable separately explained variation in the eDNA data in each of the taxonomic groups. SST and chlorophyll *a* concentration showed surfaces across nMDS plots for all markers (Fig. 2) that were simple, with gradients that were consistent across ecoregions. In contrast, human impact scores showed more complex surfaces with multiple peaks across ecoregions.

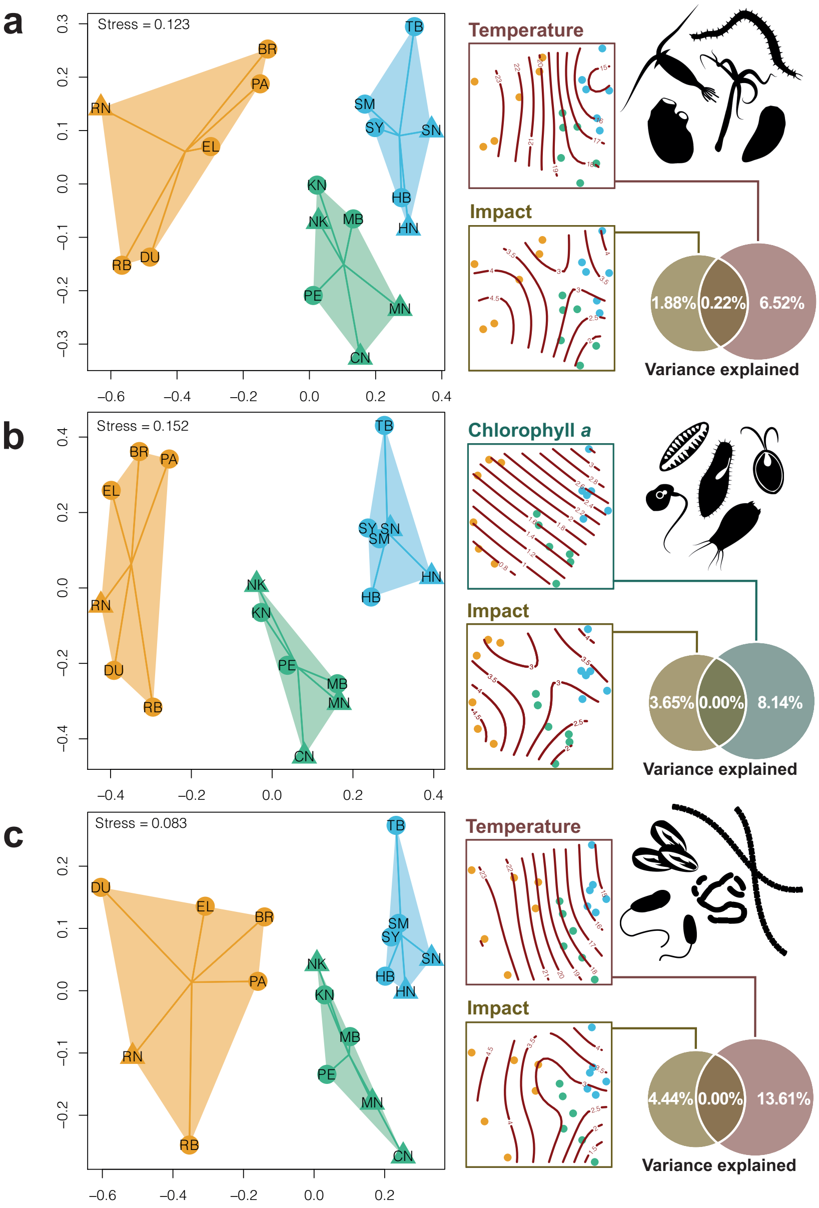


Figure 2. Observed patterns of β-diversity from environmental DNA metabarcoding of: **a** metazoans, **b** protists and **c** bacteria; based on Jaccard dissimilarities between amplicon sequence variants along the coast of South Africa. The first column of plots shows non-metric multidimensional scaling (nMDS) ordinations. Coloured hulls show the spread of the data and lines indicate the spread around the centroid grouped by coast with the east, south and west coasts denoted by orange, green and blue respectively. Site name abbreviations as in Supplementary Table 9, natural sites are denoted with triangles and artificial sites with filled circles. The second column of plots shows the same nMDS ordinations as the first column including the output of a generalised additive model with a 2D smoothed function for each of the significant environmental / impact variables overlaid; temperature – mean sea surface temperature (°C); Chlorophyll *a* – chlorophyll *a* concentration (mg m^{-3}); impact – human marine impact score (unitless measurement, see details in text) against the two nMDS axes. The Venn diagrams indicate the percentage total of the variance in the community dissimilarity explained by each significant variable, derived using variance partitioning of a distance-based redundancy analysis.

3.4.4 Distance-decay

Distance-decay slopes for all observations showed an exponential decrease in compositional similarity as the distance between sites increased (Fig. 3a). Regression models of log10 transformed compositional similarity indicating that this slope was statistically significant in all cases ($p<0.001$ for all taxonomic groups, full model output in Supplementary Note 4). Comparisons between artificial sites (e.g. recreational marinas, harbours) and natural sites (e.g. relatively unaltered rocky shore and natural harbours) in the metazoan dataset showed a significant difference in the slope between artificial and natural site comparisons ($F_{3,76}=47.73$, $p < 0.001$). No statistically significant differences were found between site types in the protist or bacteria data, and the same pattern was observed within taxonomic groups for the metazoan and protist datasets across both the COI and 18S data (Supplementary Note 4).

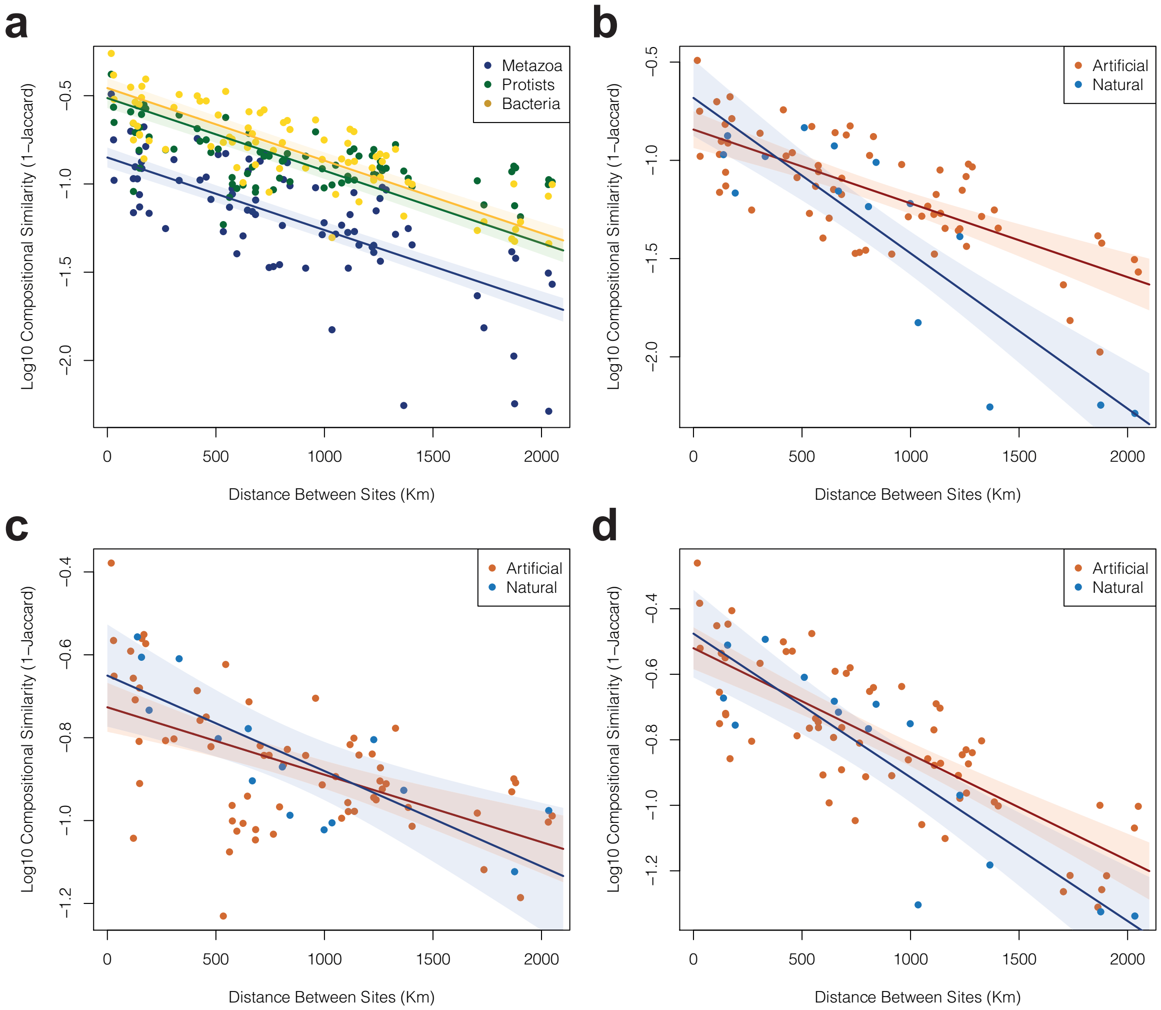


Figure 3. Plots showing distance between sites and community similarity measured using environmental DNA metabarcoding across South Africa. Logarithmically (base 10) transformed compositional similarity against distance is shown in **a**, which includes all datasets. Comparisons between artificial (coloured in red) and natural (coloured blue) sites are shown for **b** metazoans, **c** protists and **d** bacteria. 95% confidence intervals from the regression models are shown as light shaded areas around each regression slope.

3.5 Discussion

Here, we showed that metazoans, protists and bacteria have similar biogeographic patterns along an extensive and heterogenous coastline. We found that these remarkably consistent patterns could be partially explained by measured environmental conditions (chlorophyll *a* and temperature), and to a lesser extent, cumulative human impacts. Additionally, we found evidence for anthropogenically driven homogenisation of communities, but this effect was restricted to metazoans. Collectively, we provide evidence of congruent biogeographic boundaries across vastly different forms of life, and demonstrate that underlying processes, such as anthropogenic alterations, affect biogeographic patterns differentially across taxa.

Prokaryotes and eukaryotes diverged billions of years ago and have since evolved to inhabit a vast range of ecological niches. Previous studies have shown both similar (Astorga et al. 2012, Rapacciuolo et al. 2019) and dissimilar (Wang et al. 2012) patterns of β diversity between macro- and microscopic species across environmental and geographic gradients. Recent work has explored biogeographic regionalisation in marine plankton across kingdoms (Richter et al. 2019), showing that smaller planktonic organisms such as bacteria may have greater biogeographic structuring compared to larger metazoans or protists (Richter et al. 2019). Together this evidence suggests that different ecological processes drive a number of taxon-specific responses to produce patterns that are not universal across ecosystems at different spatial scales (Shade et al. 2018). Here we observed similar biogeographic patterns across life's kingdoms (Fig. 2), providing clear evidence of cross-phyla biogeographical congruence.

Our analyses suggest that environmental variables such as temperature or chlorophyll *a* concentration influenced the structure of marine communities across the study region (Fig. 2). Global studies of biogeographic patterns have shown a central role of temperature in the structuring of both microbial (Sunagawa et al. 2015) and larger planktonic life (Tittensor et al. 2010, Richter et al. 2019) across the oceans. There is growing evidence that the range boundaries of marine organisms closely track their thermal limits (Sunday et al. 2012). Therefore, a general expectation was that species would remain within their thermal niche resulting in temperature-structured communities as observed here. In contrast to temperature, salinity had a minor role in structuring the studied communities (Fig. 2), an observation previously reported in a global marine analysis (Sunagawa et al. 2015), with exceptions found in microbial (Herlemann et al. 2011) and meiofaunal (Broman et al. 2019) life in regions with unusually strong salinity gradients (e.g. Baltic Sea). The SSS range across our study system was very narrow (35.0 - 35.4ppt) and so the negligible observed effect was unsurprising. In the case of protists, biogeographic patterns showed a stronger association with primary productivity (measured here as chlorophyll *a* concentration). Previous research has shown little or no role of productivity in driving coastal and oceanic scale biodiversity patterns (Tittensor et al. 2010, Richter et al. 2019). However, these studies explored the global role of various environmental variables; the significance of more localised oceanographic systems such as upwelling (as along the western coast of southern Africa) might not be as apparent in global analyses.

Anthropogenic activities are known to alter both the physico-chemical properties of the marine environment and the trophic and ecological properties of ecosystems (Shochat et al. 2006). In our study system, human impact provided some explanatory power to understand the observed community structure, but to a much lesser extent compared to environmental variables (Fig. 2). The human impact index used here (Halpern et al. 2019) covered a large number of different types of impact (e.g. pollution, shipping intensity) but even this aggregated approach adequately explained a small proportion of the total variation in ASVs observed among sampling sites. Previous work on marine metazoans has shown a strong effect of proximate urbanisation (Kelly et al. 2016) and the ecological drivers produced through anthropogenic activities are well documented (Shochat et al. 2006). Interestingly, the pervasive and conspicuous urbanisation of the marine environment in the study area showed a much weaker effect on biogeographic patterns than other explanatory variables (Fig. 2). Anthropogenic pressures have become a major ecological driver only relatively recently in evolutionary time, with the most dramatic changes in biodiversity occurring within the 21st century (Pecl et al. 2017). It is clear that human activities are altering evolutionary trajectories (Shochat et al. 2006), either through extinction, range expansions or contractions. However, our data suggests that centuries of human impacts in our study system have not yet demonstrably altered the main observed biogeographic boundaries across taxa.

Previous work on biotic homogenisation has shown a dramatic effect on whole communities at both regional (Blouin et al. 2019, Finderup Nielsen et al. 2019) and global scales (Capinha et al. 2015, Bernardo-Madrid et al. 2019). Here, we found support for biotic homogenisation along the South African coastline only in metazoan species, with a difference in the slope of a distance-decay relationship between artificial and natural sites. This pattern was consistent for metazoans across the gene regions considered (Supplementary Note 4). Pervasive vessel activity in the region (Kaluza et al. 2010), along with evidence that artificial environments are hotspots for biological invasions (Holman et al. 2019), suggest that introduced metazoans are contributing to homogenisation of coastal communities. Further work should incorporate time series data to explore biotic homogenisation, given the significant but minor role of human impact in structuring ecological communities across the region.

Both environmental parameters and species interactions have a clear-cut effect on marine community structure across kingdoms of life (Lima-Mendez et al. 2015), but the comparative role of deterministic (environmental filtering, niche processes, etc.) and stochastic processes (ecological drift, random dispersal, etc.) in explaining the observed patterns remains uncertain. The classical deterministic theory of microbial biogeography (Baas Becking hypothesis (Baas-Becking 1934), often summarised as 'everything is everywhere but the environment selects') postulates that due to vast population sizes and dispersal, microbes are found in all environments and the variation in abiotic conditions selects for those that make up the vast majority of species in each region. This theory ignores neutral processes which have been shown to have a critical role in structuring microbial biogeography across biomes (Hanson et al. 2012). In line with previous efforts studying deterministic and stochastic processes across taxonomic kingdoms (Farjalla et al. 2012, Wu et al. 2018), we found that the majority of the observed variation could not be fully explained for both prokaryotic and eukaryotic species (Fig. 2). Indeed, recent biogeographic research in the oceans has provided both theoretical (Hellweger et al. 2014) and empirical (Richter et al. 2019) evidence of strong biogeographic patterns driven by both stochastic and deterministic forces, but much of the observed variation between communities remains unexplained. Understanding the comparative roles of different community structuring processes requires a more comprehensive examination of the observed variance between communities, species interactions, and the broader role of the environmental conditions where they live.

Several recent innovations will provide valuable data to help uncover the unexplained variation in community structure. For example, the extraction and analysis of sedimentary ancient DNA allows the reconstruction of high-resolution biodiversity change over time (Balint et al. 2018), providing evidence to evaluate the role of deterministic processes relative to temporal changes in environmental conditions. In addition, the analysis of co-occurrence networks from molecular data can provide species interaction hypotheses (e.g. Djurhuus et al. (2020)) that could be used to explore how ecological interactions structure biogeographic patterns. Finally, we expect that very high-resolution multi-spectral remote sensing data (e.g. WorldView-3, $<100m^2$) will provide unparalleled insights into the role of environmental forces structuring the distribution of ecological communities (He et al. 2015).

3.6 Methods

3.6.1 Field sampling

We sampled a range of sites along 2,000 km of coastline (Fig. 1) between October and November of 2017 (see details in Supplementary Table 9), covering the three major marine coastal ecoregions of South Africa. In order to assess the effects of anthropogenic impacts, we compared human altered 'artificial' sites (e.g. recreational marinas, harbours) and 'natural' sites (relatively unaltered rocky shore and natural harbour sites) (see Supplementary Table 9). The artificial sites were previously surveyed for marine invertebrate biodiversity by Rius et al. (2014), and six adjacent natural sites were selected for this study. The natural sites were the nearest non-developed sites with matching aspect and exposure (Fig. 1) to each of the artificial sampling sites. Three 400 ml seawater samples were filtered with 0.22 μm polyethersulfone membrane Sterivex filters (Merck Millipore, MA, USA) following the sampling scheme of Holman et al. (2019) at each sampling site. Consequently, we sampled a total of 1,200 ml of seawater per site, a volume that has been shown to differentiate fine scale (<1 km^2) community structure in marine systems (Kelly et al. 2016, Holman et al. 2019). Filters were immediately preserved at ambient temperature with the addition of 1.5 ml of Longmire's Solution for preservation until DNA extraction. Field control filters and equipment cleaning blanks were taken, transported, stored and sequenced as the rest of the field samples.

3.6.2 Environmental DNA extraction

We used a PCR-free laboratory separated from the main molecular biology laboratory facilities. No post-PCR or high concentration DNA samples were permitted in the laboratory. All surfaces and lab equipment were cleaned thoroughly before use with 1.25% sodium hypochlorite solution (3:1 dilution of household bleach). DNA extraction followed the SX$^{\text{CAPSULE}}$ method from Spens et al. (2017). Briefly, filters were first externally cleaned with sterile water and Longmire's Solution was removed from the filter outlet using a sterile syringe, 720 μl Buffer ATL (Qiagen, Hilden, Germany) and 80 μl Proteinase K (20mg/ml) was added and filters were incubated overnight at 56°C. The lysate was then removed from the filter inlet and subjected to DNA extraction using the Qiagen DNeasy Blood and Tissue Kit under the manufacturers recommended protocol. DNA was eluted using 200 µl Qiagen Buffer AE and re-eluted once to increase DNA yield. All DNA samples were checked for PCR inhibition using the Primer Design Internal Positive Control qPCR Kit (Primer Design, Southampton, UK) with 10 µl reactions under the manufacturer recommended protocol. Inhibition was detected by an increase of >1.0 Ct in reactions containing eDNA compared to reactions with extraction controls. As inhibition was detected in a minority of samples, all samples were treated using the Zymo OneStep PCR Inhibition Removal Kit (Zymo Research, California, USA) following the manufacturer recommended protocol. Inhibited samples showed no evidence for inhibition post cleaning.

3.6.3 High throughput eDNA amplicon sequencing

Different sets of primers were used to generate three separate eDNA metabarcoding libraries for all samples. Two gene regions were selected to target broad metazoan/eukaryotic diversity: a 313bp region of COI (Leray et al. 2013) and a variable length region of 18S (Zhan et al. 2013). A 16S gene region of variable length was used to target the prokaryotes (Takahashi et al. 2014). Illumina unique double-indexed metabarcoding amplicon libraries were constructed with a two-step PCR protocol as detailed in Holman et al. (2019). The first PCR setup was performed in a PCR-free laboratory. The three eDNA samples per site were pooled and three independent technical replicates were sequenced per pool. The process per sequenced pool was as follows. The first PCR reaction was conducted in triplicate in a total reaction volume of 20 µl. Each reaction contained 10 µl Amplitaq GOLD 360 2X Mastermix (Applied Biosystems, California, USA), 0.8 µl (5 nmol ml^{-1}) of each forward and reverse primers and 2 µl of undiluted environmental DNA template. The reaction conditions for PCR were an initial denaturation step at 95°C for 10 minutes followed by 20 cycles of 95°C for 30 seconds, variable annealing temp (46°C for COI, 50°C for 18S and 55°C for 16S) for 30 seconds, and extension at 72°C for 1 minute. A final extension at 72°C was performed for 10 minutes. The triplicate first PCR replicates were then pooled and cleaned using AMPure XP beads (Beckman Coulter, California, USA) at 0.8 beads:sample volume ratio following manufacturer's instructions. The second PCR reaction was conducted in a total volume of 20 µl containing 10 µl Amplitaq GOLD 360 2X Mastermix, 0.5 µl (10 nmol ml^{-1}) of both forward and reverse primers and 5 µl of undiluted cleaned PCR product from the first reaction. PCR conditions were an initial denaturation step at 95°C for 10 minutes followed by 15 cycles of 95°C for 30 seconds, annealing at 55°C for 30 seconds, and extension at 72°C for 1 minute. A final extension at 72°C was performed for 10 minutes. PCR 2 products were cleaned using AMPure XP beads as above. Negative control samples for the filters, extraction kit, PCR1 and 2 were included in library building and sequenced alongside experimental samples. Products were quantified following the manufacturer's instructions using the NEBNext Library Quant qPCR kit (New England Biolabs, Massachusetts, USA) and then normalised and pooled at an equimolar concentration for each marker. Each gene region was sequenced independently using a V3 paired-end 300bp reagent kit on the Illumina MiSeq Instrument with 5% PhiX genomic library added to increase sequence diversity.

3.6.4 Bioinformatics

Raw sequences were de-multiplexed using the GenerateFastQ (v2.0.0.9) module on the MiSeq control software (v3.0.0.105). Cutadapt (v2.3) (Martin 2011) was used to filter sequences to include only those that contained both the forward and reverse primer sequence across both read pairs for each gene fragment, remaining sequences then had the primer region removed for each gene fragment using the default settings. Sequences were denoised using the DADA2 pipeline (v1.12) (Callahan et al. 2016) in R (v3.6.1) (R_Core_Team 2019) with the default parameters unless noted as follows. Sequences were filtered to retain only pairs of reads with an expected error of 1 or fewer per read. Read trimming was performed after manual examination of the read quality profile, the forward reads were trimmed to 250bp (COI), 240bp (18S) and 240 bp (16S) and the reverse reads were trimmed to 230bp (COI), 220bp (18S) and 220 bp (16S). As each marker was sequenced separately, the differences in read trimming length reflect typical variation in sequencing runs rather than any biological difference. The error rates per run were estimated and used to perform the denoising using the DADA2 algorithm. The denoised sequence pairs were then merged and resulting sequences were truncated if they were outside of the expected gene fragment range (303-323bp for COI, 400-450bp for 18S and 390-450bp for 16S). Chimeras were identified and removed before assembling a sample by ASV table for analysis. The denoised ASVs were then curated using the default settings of the LULU algorithm (Frøslev et al. 2017) which merges sequences based on sequence similarity and co-occurrence. Assigning taxonomy to a set of unknown sequences is a difficult task, particularly considering many marine species lack DNA barcodes, are undescribed, or have erroneous barcodes in online public databases. We therefore focused our analysis at a higher taxonomic level than species, assigning taxonomy to sequences from the COI and 18S data as follows. The RDP classifier (v2.13) (Wang et al. 2007) was used to assign taxonomy for COI using a previously published COI database (Porter and Hajibabaei 2018) (v4.0) and a modified version of the SILVA database (Quast et al. 2013) (v3.2 from https://github.com/terrimporter/18SClassifier). As species level assignments have been shown to be accurate for COI data (Holman et al. 2019) an unconstrained (no limits on sequence similarity or match length) BLAST search (v2.6.0+) was performed for each sequence against the entire National Centre for Biotechnology Information *nt* database (downloaded on 16th May 2019), 200 hits per sequence were retained (-*num_alignments*). These sequences were then parsed using an R script to exclude hits below 65% coverage, remaining assignments with percent identity above 97% for COI were used to collapse reads for ASVs assigned to the same species. Recent analyses have suggested that only exact (100% identity) matching of sequences to reference data is appropriate for species assignment for the prokaryotic 16S region (Edgar 2018). The 16S sequences were matched to the SILVA database (release 132) (Quast et al. 2013) using the default settings of the *assignTaxonomy* function from the DADA2 package to assign taxonomy at genus level or above. The incidence of NUMTs (nuclear mitochondrial DNA) and chimeras in the final ASV list was evaluated following Supplementary Note 5.

The following quality control filters were applied to the ASV by sample table produced by DADA2. First, the minimum number of reads per observation was set at three. Any ASVs not represented in at least one other sample were discarded. ASVs were then filtered to retain only those found in all three technical replicates. For any ASV found in the negative control samples, the largest value among the read count across all negative control samples was used as the zero value for all other samples (i.e. any smaller values found in non-control samples were set to zero). The COI and 18S datasets were then subset by the RDP classifier taxonomic assignments to produce datasets for the protists and metazoans as follows. Phylum level assignments above a threshold of 30, a value well above that shown to accurately assign phylum level taxonomy (Porter and Hajibabaei 2018), were parsed to include phyla that contained only metazoan or protist members for each group respectively, other assignments or unknown assignments were discarded. This resulted in a protist and metazoan dataset for each marker, which were subsequently used for separate analyses using these groupings. The 16S data was parsed to include only bacterial ASVs. Within each taxonomic dataset samples were then rarefied to the smallest number of reads (see Supplementary Table 1). Technical replicates were then collapsed to produce a dataset containing the mean value of rarefied reads per ASV. Finally, ASVs assigned using BLAST (with no cases of multiple matches of equal quality) to the same species in the COI dataset were combined by summing reads per site. The taxonomic assignment method used for the 16S data assigns to genus, and species level assignments are not possible for all taxa using the selected 18S region, so no ASVs from these datasets were collapsed. In order to explore broad scale patterns of taxonomic diversity, the number of ASVs per phyla and number of rarefied reads per phyla were collapsed to produce per site assessments of taxonomic composition. As phylum level phylogeny is not resolved for all protist species, the protist dataset was grouped by supergroup designations according to Burki et al. (2020). For plots, phyla represented by less than 2% of ASV counts were concatenated in an ‘other’ category.

3.6.5 Environmental, human impact and geographic data

In situ environmental data reflects a snapshot of the total conditions experienced across the lifetime of the species that make up marine communities. Therefore, abiotic variables for the sites covering an ecologically relevant timescale were sourced as follows. High resolution (1 km^2) remote sensing average daily sea surface temperature data derived from multiple satellite deployments, combined with *in situ* data (JPL_OurOceanProject 2010) was parsed in R to find the nearest datapoint to each site. For each point, a mean from two years of data from November 2017 was calculated. Interpolated average (2005-2017) sea surface salinity data (0.25° grid resolution) generated using gliders, oceanographic casts etc. from the 2018 World Ocean Atlas (Zweng 2019) was parsed to include only surface data for the sites. Monthly global ocean colour data (4 km^2) derived from multiple satellite deployments (European_Space_Agency 2020) was parsed to calculate an average value for chlorophyll *a* density per site across two years from November 2017. Finally, a previously described (Halpern et al. 2019) 1 km^2 global resolution cumulative index for anthropogenic impact on marine ecosystems, comprising fishing pressure, climate change, shipping and land-based pollution, was parsed to produce a value for each site cumulatively across the entire period for which data were available (2003-2013). These global datasets have excellent temporal resolution, but are only appropriate for testing large-scale patterns as they have limited ability to discriminate highly localised observations.

3.6.6 Ecological statistics

Analyses were conducted in R (v3.6.1) unless otherwise stated. Differences in the mean number of ASVs per coastline were assessed using an ANOVA after testing for normally distributed residuals using a Shapiro-Wilk test and equal variance between coasts using a Bartlett test. A Tukey's Honest Significant difference test was used to evaluate significant ANOVA results. Differences in community similarity were assessed using a PERMANOVA (Anderson 2014) implemented with the function *adonis* from the package *vegan* (v2.5-6) (Oksanen et al. 2011) to assess differences in multivariate centroids and dispersion between coastlines. The PERMANOVA was conducted on a matrix of Jaccard dissimilarities as this ecological index has been shown to be appropriate for biogeographical studies (Kreft and Jetz 2010). Significant pairwise differences were assessed using the function *adonis.pair* from the *EcolUtils* package (v0.1) (Salazar 2018). To analyse if groups of samples have a difference in intra-group community variation, also known as heterogeneity of multivariate dispersion, the PERMDISP2 procedure(Anderson 2006) was used, implemented in the function *betadisper* from the *vegan* package. The pairwise group differences in heterogeneity of multivariate dispersion in the case of a global significant result from *betadisper* were analysed using a Tukey's Honest Significant difference test. Non-metric multidimensional scaling ordinations (nMDS) were calculated using Jaccard dissimilarities and the function *metaMDS* from the *vegan* package.

The influence of the abiotic and human impact data on the observed patterns of beta diversity were evaluated as follows. It has previously been common to use a partial Mantel test to evaluate the effect of a distance matrix (frequently environmental variables) on a second distance matrix (species composition) while 'cancelling out' the effect of a third matrix (geographic distance). However, this approach has been shown to be sensitive to spatial autocorrelation common in ecological datasets (Crabot et al. 2019). A recently developed method (Crabot et al. 2019), which corrects spurious inflations of the parameter estimate for Mantel tests, was implemented. Across each taxonomic group Mantel tests were conducted comparing Jaccard dissimilarity against Euclidean distance for each environmental variable. For each test Moran spectral randomisation was performed including the geographic distance data with 10,000 permutations to assess statistical significance using the *msr* function from the *adespatial* package (v0.3-8). Partial Mantel tests were conducted with 10,000 permutations using the *mantel.partial* function from the *vegan* package for comparison.

Explanatory variables which had some correlation with the community dissimilarity after adjusting for geographic distance were then evaluated as follows. First, a distance-based redundancy analysis (dbRDA) (McArdle and Anderson 2001), regressing site Jaccard dissimilarities against all remaining variables, was performed using the function *dbrda* from the *vegan* package. The significance of terms was assessed with 10,000 permutations. The dbRDA ordination allows us to examine linear changes in the beta diversity in response to a number of predictor variables in tandem, and also to explore their relative impact. The function *varpart* from the *vegan* package was then used to partition the variance in the community dissimilarity by the environmental variables. We then used a generalised additive model to visualise the variation of each significant variable across the nMDS space via a restricted maximum likelihood 2D smoother, implemented in the function *ordisurf* from the *vegan* package.

Distance-decay relationships were explored by first measuring compositional similarity (1-Jaccard index) for each pair of sites, and then calculating distances between pairs of sites by drawing a continuous transect 1 km offshore parallel to the high-water mark using Google Earth Pro (v7.3.2.5776), taking the distance along the transect to measure distance between sites. We then used these data in least-square regression models using the function *lm* with an interaction function between distance and site type (artificial or natural) terms against compositional similarity as a response term. The compositional similarity values were log10 transformed to linearise the response, untransformed values of zero (no overlap of species) were omitted to avoid infinite response variable values.

3.7 References

Anderson, M. J. 2006. Distance-Based Tests for Homogeneity of Multivariate Dispersions. Biometrics **62**:245-253.

Anderson, M. J. 2014. Permutational multivariate analysis of variance (PERMANOVA). Pages 1-15 *in* N. Balakrishnan, T. Colton, B. Everitt, W. Piegorsch, F. Ruggeri, and J. L. Teugels, editors. Wiley Stats Ref: Statistics Reference Online. John Wiley & Sons, Ltd, Chichester, UK.

Astorga, A., J. Oksanen, M. Luoto, J. Soininen, R. Virtanen, and T. Muotka. 2012. Distance decay of similarity in freshwater communities: do macro- and microorganisms follow the same rules? Global Ecology and Biogeography **21**:365-375.

Awad, A. A., C. L. Griffiths, and J. K. Turpie. 2002. Distribution of South African marine benthic invertebrates applied to the selection of priority conservation areas. Diversity and Distributions **8**:129-145.

Azam, F., and F. Malfatti. 2007. Microbial structuring of marine ecosystems. Nature Reviews Microbiology **5**:782-791.

Baas-Becking, L. G. M. 1934. Geobiologie; of inleiding tot de milieukunde. WP Van Stockum & Zoon NV.

Balint, M., M. Pfenninger, H. P. Grossart, P. Taberlet, M. Vellend, M. A. Leibold, G. Englund, and D. Bowler. 2018. Environmental DNA time series in ecology. Trends in Ecology and Evolution **33**:945-957.

Beck, J., L. Ballesteros-Mejia, C. M. Buchmann, J. Dengler, S. A. Fritz, B. Gruber, C. Hof, F. Jansen, S. Knapp, H. Kreft, A.-K. Schneider, M. Winter, and C. F. Dormann. 2012. What's on the horizon for macroecology? Ecography **35**:673-683.

Bernardo-Madrid, R., J. Calatayud, M. González-Suárez, M. Rosvall, P. M. Lucas, M. Rueda, A. Antonelli, and E. Revilla. 2019. Human activity is altering the world's zoogeographical regions. Ecology Letters **22**:1297-1305.

Blouin, D., S. Pellerin, and M. Poulin. 2019. Increase in non-native species richness leads to biotic homogenization in vacant lots of a highly urbanized landscape. Urban Ecosystems **22**:879-892.

Blowes, S. A., S. R. Supp, L. H. Antao, A. Bates, H. Bruelheide, J. M. Chase, F. Moyes, A. Magurran, B. McGill, I. H. Myers-Smith, M. Winter, A. D. Bjorkman, D. E. Bowler, J. E. K. Byrnes, A. Gonzalez, J. Hines, F. Isbell, H. P. Jones, L. M. Navarro, P. L. Thompson, M. Vellend, C. Waldock, and M. Dornelas. 2019. The geography of biodiversity change in marine and terrestrial assemblages. Science **366**:339-345.

Broman, E., C. Raymond, C. Sommer, J. S. Gunnarsson, S. Creer, and F. J. A. Nascimento. 2019. Salinity drives meiofaunal community structure dynamics across the Baltic ecosystem. Molecular Ecology **28**:3813-3829.

Burki, F., A. J. Roger, M. W. Brown, and A. G. Simpson. 2020. The new tree of eukaryotes. Trends in Ecology & Evolution **35**:43-55.

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. Johnson, and S. P. Holmes. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods **13**:581-583.

Capinha, C., F. Essl, H. Seebens, D. Moser, and H. M. Pereira. 2015. The dispersal of alien species redefines biogeography in the Anthropocene. Science **348**:1248-1251.

Costello, M. J., P. Tsai, P. S. Wong, A. K. L. Cheung, Z. Basher, and C. Chaudhary. 2017. Marine biogeographic realms and species endemicity. Nature Communications **8**.

Crabot, J., S. Clappe, S. Dray, and T. Datry. 2019. Testing the Mantel statistic with a spatially-constrained permutation procedure. Methods in Ecology and Evolution **10**:532-540.

Deiner, K., H. M. Bik, E. Machler, M. Seymour, A. Lacoursiere-Roussel, F. Altermatt, S. Creer, I. Bista, D. M. Lodge, N. de Vere, M. E. Pfrender, and L. Bernatchez. 2017. Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. Molecular Ecology **26**:5872-5895.

Djurhuus, A., C. J. Closek, R. P. Kelly, K. J. Pitz, R. P. Michisaki, H. A. Starks, K. R. Walz, E. A. Andruszkiewicz, E. Olesin, K. Hubbard, E. Montes, D. Otis, F. E. Muller-Karger, F. P. Chavez, A. B. Boehm, and M. Breitbart. 2020. Environmental DNA reveals seasonal shifts and potential interactions in a marine community. Nature Communications **11**:254.

Dornelas, M., N. J. Gotelli, B. McGill, H. Shimadzu, F. Moyes, C. Sievers, and A. E. Magurran. 2014. Assemblage time series reveal biodiversity change but not systematic loss. Science **344**:296-299.

Edgar, R. C. 2018. Updating the 97% identity threshold for 16S ribosomal RNA OTUs. Bioinformatics **34**:2371-2375.

Emanuel, B. P., R. H. Bustamante, G. M. Branch, S. Eekhout, and F. J. Odendaal. 1992. A zoogeographic and functional approach to the selection of marine reserves on the west coast of South Africa. South African Journal of Marine Science **12**:341-354.

European_Space_Agency. 2020. Ocean Colour Climate Change Initiative Dataset, Version 4.2.

Farjalla, V. F., D. S. Srivastava, N. A. C. Marino, F. D. Azevedo, V. Dib, P. M. Lopes, A. S. Rosado, R. L. Bozelli, and F. A. Esteves. 2012. Ecological determinism increases with organism size. Ecology **93**:1752-1759.

Ficetola, G. F., F. Mazel, and W. Thuiller. 2017. Global determinants of zoogeographical boundaries. Nature Ecology and Evolution **1**:89.

Finderup Nielsen, T., K. Sand-Jensen, M. Dornelas, and H. H. Bruun. 2019. More is less: net gain in species richness, but biotic homogenization over 140 years. Ecology Letters **22**:1650-1657.

Frøslev, T. G., R. Kjøller, H. H. Bruun, R. Ejrnæs, A. K. Brunbjerg, C. Pietroni, and A. J. Hansen. 2017. Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. Nature Communications **8**:1188.

Griffiths, C. L., T. B. Robinson, L. Lange, and A. Mead. 2010. Marine biodiversity in South Africa: an evaluation of current states of knowledge. PLoS One 5:e12008.

Griffiths, C. L., L. Van Sittert, P. B. Best, A. C. Brown, B. M. Clark, P. A. Cook, R. J. M. Crawford, J. H. M. David, B. Davies, M. H. Griffiths, K. Hutchings, A. Jerardino, N. Kruger, S. Lamberth, R. W. Leslie, R. Melville-Smith, R. Tarr, and C. D. Van der Lingen. 2004. Impacts of human activities on marine animal life in the Benguela: A historical overview. Oceanography and Marine Biology: an Annual Review 42:303-392.

Halpern, B. S., M. Frazier, J. Afflerbach, J. S. Lowndes, F. Micheli, C. O'Hara, C. Scarborough, and K. A. Selkoe. 2019. Recent pace of change in human impact on the world's ocean. Scientific Reports 9:11609.

Hanson, C. A., J. A. Fuhrman, M. C. Horner-Devine, and J. B. Martiny. 2012. Beyond biogeographic patterns: processes shaping the microbial landscape. Nature Reviews Microbiology 10:497-506.

He, K. S., B. A. Bradley, A. F. Cord, D. Rocchini, M. N. Tuanmu, S. Schmidtlein, W. Turner, M. Wegmann, and N. Pettorelli. 2015. Will remote sensing shape the next generation of species distribution models? Remote Sensing in Ecology and Conservation 1:4-18.

Hellweger, F. L., E. van Sebille, and N. D. Fredrick. 2014. Biogeographic patterns in ocean microbes emerge in a neutral agent-based model. Science 345:1346-1349.

Herlemann, D. P., M. Labrenz, K. Jurgens, S. Bertilsson, J. J. Waniek, and A. F. Andersson. 2011. Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. ISME Journal 5:1571-1579.

Holman, L. E., M. de Bruyn, S. Creer, G. Carvalho, J. Robidart, and M. Rius. 2019. Detection of introduced and resident marine species using environmental DNA metabarcoding of sediment and water. Scientific Reports 9:11559.

Holt, B. G., J. P. Lessard, M. K. Borregaard, S. A. Fritz, M. B. Araujo, D. Dimitrov, P. H. Fabre, C. H. Graham, G. R. Graves, K. A. Jonsson, D. Nogues-Bravo, Z. Wang, R. J. Whittaker, J. Fjeldsa, and C. Rahbek. 2013. An update of Wallace's zoogeographic regions of the world. Science 339:74-78.

JPL_OurOceanProject. 2010. GHRSST Level 4 G1SST Global Foundation Sea Surface Temperature Analysis.

Kaluza, P., A. Kolzsch, M. T. Gastner, and B. Blasius. 2010. The complex network of global cargo ship movements. Journal of The Royal Society Interface 7:1093-1103.

Kelly, R. P., J. L. O'Donnell, N. C. Lowell, A. O. Shelton, J. F. Samhouri, S. M. Hennessey, B. E. Feist, and G. D. Williams. 2016. Genetic signatures of ecological diversity along an urbanization gradient. PeerJ 4:e2444.

Kocsis, A. T., C. J. Reddin, and W. Kiessling. 2018. The stability of coastal benthic biogeography over the last 10 million years. Global Ecology and Biogeography 27:1106-1120.

Kreft, H., and W. Jetz. 2010. A framework for delineating biogeographical regions based on species distributions. Journal of Biogeography 37:2029-2053.

Leray, M., J. Y. Yang, C. P. Meyer, S. C. Mills, N. Agudelo, V. Ranwez, J. T. Boehm, and R. J. Machida. 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.

Lima-Mendez, G., K. Faust, N. Henry, J. Decelle, S. Colin, F. Carcillo, S. Chaffron, J. C. Ignacio-Espinosa, S. Roux, F. Vincent, L. Bittner, Y. Darzi, J. Wang, S. Audic, L. Berline, G. Bontempi, A. M. Cabello, L. Coppola, F. M. Cornejo-Castillo, F. d'Ovidio, L. De Meester, I. Ferrera, M. J. Garet-Delmas, L. Guidi, E. Lara, S. Pesant, M. Royo-Llonch, G. Salazar, P. Sanchez, M. Sebastian, C. Souffreau, C. Dimier, M. Picheral, S. Searson, S. Kandels-Lewis, c. Tara Oceans, G. Gorsky, F. Not, H. Ogata, S. Speich, L. Stemmann, J. Weissenbach, P. Wincker, S. G. Acinas, S. Sunagawa, P. Bork, M. B. Sullivan, E. Karsenti, C. Bowler, C. de Vargas, and J. Raes. 2015. Determinants of community structure in the global plankton interactome. Science 348:1262073.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet Journal 17:10-12.

McArdle, B. H., and M. J. Anderson. 2001. Fitting Multivariate Models To Community Data: A Comment On Distance-Based Redundancy Analysis. Ecology 82:290-297.

Mouillot, D., D. R. Bellwood, C. Baraloto, J. Chave, R. Galzin, M. Harmelin-Vivien, M. Kulbicki, S. Lavergne, S. Lavorel, N. Mouquet, C. E. Paine, J. Renaud, and W. Thuiller. 2013. Rare species support vulnerable functions in high-diversity ecosystems. PLoS Biology 11:e1001569.

Naeem, S., J. E. Duffy, and E. Zavaleta. 2012. The functions of biological diversity in an age of extinction. Science 336:1401-1406.

Oksanen, J., F. G. Blanchet, R. Kindt, P. Legendre, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, and H. Wagner. 2011. Vegan: community ecology package. R package, Version 2.5-6.

Olden, J. D., and T. P. Rooney. 2006. On defining and quantifying biotic homogenization. Global Ecology and Biogeography 15:113-120.

Pecl, G. T., M. B. Araujo, J. D. Bell, J. Blanchard, T. C. Bonebrake, I. C. Chen, T. D. Clark, R. K. Colwell, F. Danielsen, B. Evengard, L. Falconi, S. Ferrier, S. Frusher, R. A. Garcia, R. B. Griffis, A. J. Hobday, C. Janion-Scheepers, M. A. Jarzyna, S. Jennings, J. Lenoir, H. I. Linnetved, V. Y. Martin, P. C. McCormack, J. McDonald, N. J. Mitchell, T. Mustonen, J. M. Pandolfi, N. Pettorelli, E. Popova, S. A. Robinson, B. R. Scheffers, J. D. Shaw, C. J. Sorte, J. M. Strugnell, J. M. Sunday, M. N. Tuanmu, A. Verges, C. Villanueva, T. Wernberg, E. Wapstra, and S. E. Williams. 2017. Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. Science 355:6332.

Porter, T. M., and M. Hajibabaei. 2018. Automated high throughput animal CO1 metabarcode classification. Scientific Reports 8:4226.

Quast, C., E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, and F. O. Glockner. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Research **41**:D590-596.

R_Core_Team. 2019. R: a language and environment for statistical computing. ISBN 3-900051-07-0.

Rapacciuolo, G., J. M. Beman, L. M. Schiebelhut, and M. N. Dawson. 2019. Microbes and macro-invertebrates show parallel β -diversity but contrasting α -diversity patterns in a marine natural experiment. Proceedings of the Royal Society B: Biological Sciences **286**:20190999.

Richter, D. J., R. Watteaux, T. Vannier, J. Leconte, P. Frémont, G. Reygondeau, N. Maillet, N. Henry, G. Benoit, A. Fernàndez-Guerra, S. Suweis, R. Narci, C. Berney, D. Eveillard, F. Gavory, L. Guidi, K. Labadie, E. Mahieu, J. Poulain, S. Romac, S. Roux, C. Dimier, S. Kandels, M. Picheral, S. Searson, S. Pesant, J.-M. Aury, J. R. Brum, C. Lemaitre, E. Pelletier, P. Bork, S. Sunagawa, L. Karp-Boss, C. Bowler, M. B. Sullivan, E. Karsenti, M. Mariadassou, I. Probert, P. Peterlongo, P. Wincker, C. de Vargas, M. Ribera d'Alcalà, D. Iudicone, and O. Jaillon. 2019. Genomic evidence for global ocean plankton biogeography shaped by large-scale current systems. bioRxiv:867739.

Rius, M., S. Clusella-Trullas, C. D. McQuaid, R. A. Navarro, C. L. Griffiths, C. A. Matthee, S. von der Heyden, and X. Turon. 2014. Range expansions across ecoregions: interactions of climate change, physiology and genetic diversity. Global Ecology and Biogeography **23**:76-88.

Salazar, G. 2018. EcolUtils: Utilities for community ecology analysis. R package version 0.1.

Shade, A., R. R. Dunn, S. A. Blowes, P. Keil, B. J. M. Bohannan, M. Herrmann, K. Küsel, J. T. Lennon, N. J. Sanders, D. Storch, and J. Chase. 2018. Macroecology to Unite All Life, Large and Small. Trends in Ecology and Evolution **33**:731-744.

Shochat, E., P. S. Warren, S. H. Faeth, N. E. McIntyre, and D. Hope. 2006. From patterns to emerging processes in mechanistic urban ecology. Trends in Ecology and Evolution **21**:186-191.

Spalding, M. D., H. E. Fox, B. S. Halpern, M. A. McManus, J. Molnar, G. R. Allen, N. Davidson, Z. A. Jorge, A. L. Lombana, S. A. Lourie, K. D. Martin, E. McManus, J. Molnar, C. A. Recchia, and J. Robertson. 2007. Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. BioScience **57**:573-583.

Spens, J., A. R. Evans, D. Halfmaerten, S. W. Knudsen, M. E. Sengupta, S. S. T. Mak, E. E. Sigsgaard, M. Hellström, and D. Yu. 2017. Comparison of capture and storage methods for aqueous macrobial eDNA using an optimized extraction protocol: advantage of enclosed filter. Methods in Ecology and Evolution **8**:635-645.

Stuart-Smith, R. D., A. E. Bates, J. S. Lefcheck, J. E. Duffy, S. C. Baker, R. J. Thomson, J. F. Stuart-Smith, N. A. Hill, S. J. Kininmonth, L. Airoldi, M. A. Becerro, S. J. Campbell, T. P. Dawson, S. A. Navarrete, G. A. Soler, E. M. Strain, T. J. Willis, and G. J. Edgar. 2013. Integrating abundance and functional traits reveals new global hotspots of fish diversity. Nature **501**:539-542.

Sunagawa, S., L. P. Coelho, S. Chaffron, J. R. Kultima, K. Labadie, G. Salazar, B. Djahanschiri, G. Zeller, D. R. Mende, A. Alberti, F. M. Cornejo-Castillo, P. I. Costea, C. Cruaud, F. d'Ovidio, S. Engelen, I. Ferrera, J. M. Gasol, L. Guidi, F. Hildebrand, F. Kokoszka, C. Lepoivre, G. Lima-Mendez, J. Poulain, B. T. Poulos, M. Royo-Llonch, H. Sarmento, S. Vieira-Silva, C. Dimier, M. Picheral, S. Searson, S. Kandels-Lewis, c. Tara Oceans, C. Bowler, C. de Vargas, G. Gorsky, N. Grimsley, P. Hingamp, D. Iudicone, O. Jaillon, F. Not, H. Ogata, S. Pesant, S. Speich, L. Stemmann, M. B. Sullivan, J. Weissenbach, P. Wincker, E. Karsenti, J. Raes, S. G. Acinas, and P. Bork. 2015. Ocean plankton. Structure and function of the global ocean microbiome. Science **348**:1261359.

Sunday, J. M., A. E. Bates, and N. K. Dulvy. 2012. Thermal tolerance and the global redistribution of animals. Nature Climate Change **2**:686-690.

Takahashi, S., J. Tomita, K. Nishioka, T. Hisada, and M. Nishijima. 2014. Development of a prokaryotic universal primer for simultaneous analysis of Bacteria and Archaea using next-generation sequencing. PLoS One **9**:e105592.

Tilman, D., F. Isbell, and J. M. Cowles. 2014. Biodiversity and Ecosystem Functioning. Annual Review of Ecology, Evolution, and Systematics **45**:471-493.

Tittensor, D. P., C. Mora, W. Jetz, H. K. Lotze, D. Ricard, E. Vanden Berghe, and B. Worm. 2010. Global patterns and predictors of marine biodiversity across taxa. Nature **466**:1098-U1107.

Wallace, A. R. 1876. The geographical distribution of animals, with a study of the relations of living and extinct faunas as elucidating the past changes of the earth's surface. Macmillan and Co., London,.

Wang, J., J. Soininen, Y. Zhang, B. Wang, X. Yang, and J. Shen. 2012. Patterns of elevational beta diversity in micro- and macroorganisms. Global Ecology and Biogeography **21**:743-750.

Wang, Q., G. M. Garrity, J. M. Tiedje, and J. R. Cole. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and environmental microbiology **73**:5261-5267.

Wu, W. X., H. P. Lu, A. Sastri, Y. C. Yeh, G. C. Gong, W. C. Chou, and C. H. Hsieh. 2018. Contrasting the relative importance of species sorting and dispersal limitation in shaping marine bacterial versus protist communities. ISME Journal **12**:485-494.

Zaffosa, A., S. Finnegan, and S. E. Peters. 2017. Plate tectonic regulation of global marine animal diversity. Proceedings of the National Academy of Sciences of the United States of America **114**:5653-5658.

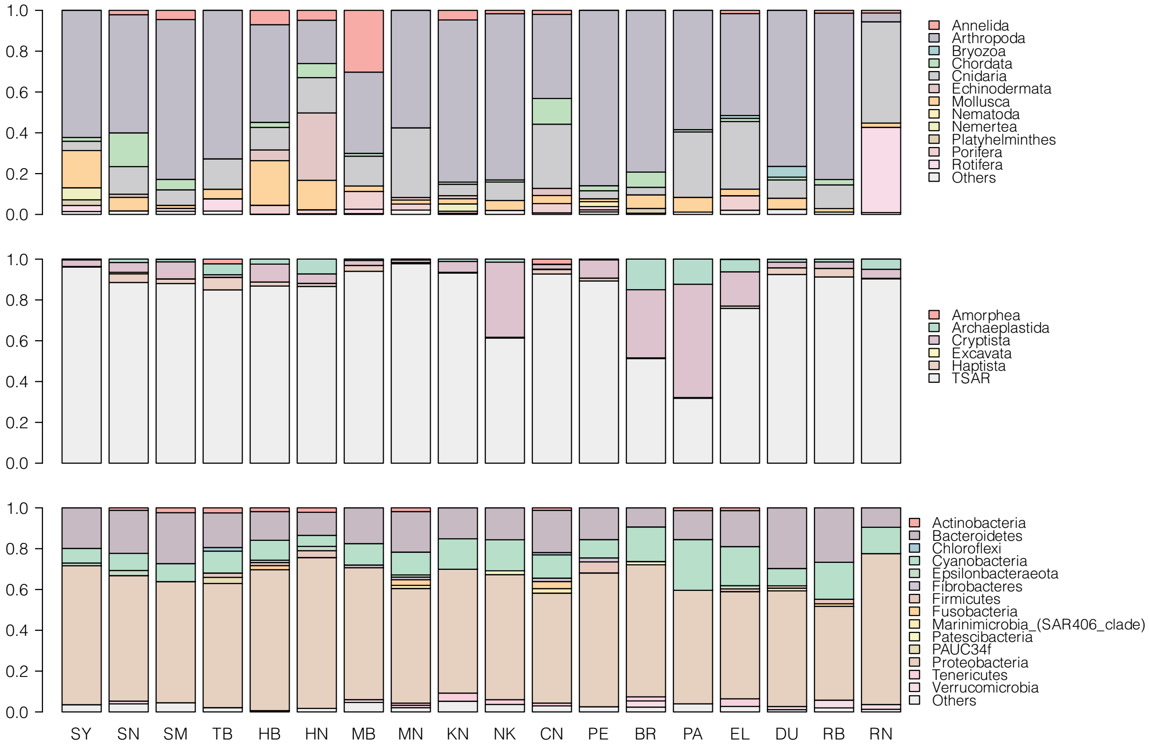
Zhan, A., M. Hulák, F. Sylvester, X. Huang, A. A. Adebayo, C. L. Abbott, S. J. Adamowicz, D. D. Heath, M. E. Cristescu, and H. J. MacIsaac. 2013. High sensitivity of 454 pyrosequencing for detection of rare species in aquatic communities. Methods in Ecology and Evolution **4**:558-565.

Zweng, M. M., J. R. Reagan, D. Seidov, T. P. Boyer, R. A. Locarnini, H. E. Garcia, A. V. Mishonov, O. K. Baranova, K. Weathers, C. R. Paver, and I. Smolyar,. 2019. World Ocean Atlas 2018, Volume 2: Salinity.50.

3.8 Supplementary Information

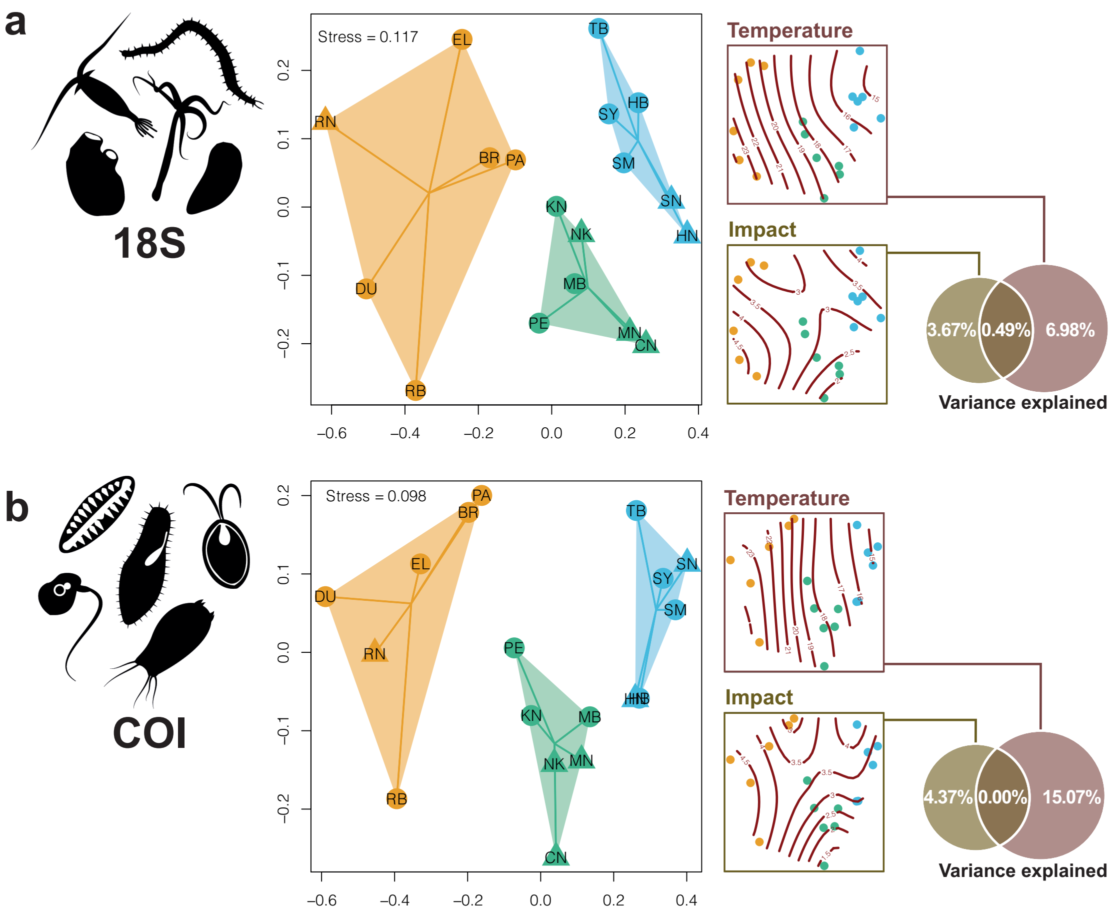
3.8.1 Supplementary Figure 1

Bar charts indicating the proportion of reads assigned per phyla (metazoans/bacteria) or supergroup (protists) from environmental DNA metabarcoding of seawater collected from sites across South Africa. The three rows correspond with data from metazoans (top), protists (middle) and bacteria (bottom). Site name abbreviations as in Supplementary Table 9.



3.8.2 Supplementary Figure 2

Observed patterns of β-diversity from environmental DNA metabarcoding of: ***a*** *metazoans from the 18S dataset and* ***b*** *protists from the COI dataset; based on Jaccard dissimilarities between amplicon sequence variants along the coast of South Africa. The first column of plots shows non-metric multidimensional scaling (nMDS) ordinations. Coloured hulls show the spread of the data and lines indicate the spread around the centroid grouped by coast with the east, south and west coasts denoted by orange, green and blue respectively. Site name abbreviations as in Supplementary Table 9. Natural sites are denoted with triangles and artificial sites with filled circles. The second column of plots shows the same nMDS ordinations as the first column including the output of a generalised additive model with a 2D smoothed function for each of the significant environmental / impact variables overlaid; temperature – mean sea surface temperature (°C); impact – human marine impact score (unitless measurement, see details in main text) against the two nMDS axes. The Venn diagram charts indicate the percentage total of variance in community dissimilarity explained by each significant variable, derived using variance partitioning of a distance-based redundancy analysis.*



3.8.3 Supplementary Table 1

Descriptive statistics of bioinformatic parameters from environmental DNA metabarcoding data of natural and artificial sites along the coast of South Africa. Each column indicates **A** *the gene fragment used in the metabarcoding experiment; COI – cytochrome c oxidase subunit I; 18S –nuclear small subunit ribosomal DNA, 16S – prokaryotic small subunit ribosomal DNA* **B** *the dataset subset into metazoan (COI), protist (18S) and bacterial (16S) datasets. The first two rows indicate the number of reads and standard deviation per sample. The third indicates the lowest number of reads per sample used to rarefy the data for analyses. The fourth and fifth rows indicate the number of Amplicon Sequence Variants (ASVs) from the dataset and the number that did not have a (70% percentage identity) match in the NCBI nt database. The final row indicates the number of ASVs with high quality taxonomic assignment; a full length hit above 97% identity for COI and 99% for 18S; and a genus level assignment using the default settings of the 'assignTaxonomy' function from the DADA2 R package.*

A

	COI	18S	16S
Raw mean reads per sample	207,675.8	361,868.6	473,062.8
Standard Deviation	112,122.6	110,399.3	106,473.1
N_{reads} Rarefaction	61,958	147,307	115,960
ASVs in analysis	4,867	2,364	2,826
ASVs with no taxonomic assignment	2,824	131	8
ASVs with high quality BLAST nt assignment	150	265	1,660

B

	Metazoa	Protists	Bacteria
N_{reads} Rarefaction	8,682	12,161	115,959
ASVs in analysis	1,054	1,433	2,826

3.8.4 Supplementary Table 2

Analysis of variance (ANOVA) testing differences in amplicon sequence variant richness between coasts for datasets by ***A*** *taxa and* ***B*** *marker, a Tukey post-hoc test for bacteria is shown.*

Significant p-values at α=0.05 level are indicated in ***bold.***

A

ANOVA	Df	SumofSquares	MeanSquare	F	P value
Metazoa					
Coast	2	4054	2027	1.941	0.178
Residuals	15	15664	1044		
Protists					
Coast	2	7819	3910	1.416	0.273
Residuals	15	41404	2760		
Bacteria					
Coast	2	79844	39922	7.178	**0.007**
Residuals	15	83422	5561		

TukeyPostHoc

Bacteria				
	diff	lwr	upr	P value
South-East	146.167	34.330	258.004	**0.012**
West-East	10.333	-101.504	122.170	0.969
West-South	-135.833	-247.670	-23.996	**0.017**

B

ANOVA	Df	SumofSquares	MeanSquare	F	P value
COI					
Coast	2	0.115	0.05763	3.64	0.051
Residuals	15	0.237	0.01583		

18S					
Coast	2	28434	14217	1.824	0.195
Residuals	15	116945	7796		
16S					
Coast	2	79844	39922	7.178	**0.007**
Residuals	15	83422	5561		

3.8.5 Supplementary Table 3

Permutational analysis of variance (PERMANOVA) model outputs based on a Jaccard index of amplicon sequence variants from environmental DNA metabarcoding of sites across South Africa. Models are presented for **A** *three taxonomic groups: metazoa, protists and bacteria.* **B** *three genes: COI – cytochrome c oxidase subunit I; 18S – nuclear small subunit ribosomal DNA; 16S – prokaryotic small subunit ribosomal DNA. Significant p-values at* α*=0.05 level are indicated in* ***bold****.*

A

	Df	SumsOfSqs	MeanSqs	Pseudo F	R^2	P value
Metazoa						
Coast	2	1.511	0.755	1.981	0.209	**0.001**
Residuals	15	5.721	0.381		0.791	
Total	17	7.232				
Protists						
Coast	2	1.601	0.800	2.656	0.262	**0.001**
Residuals	15	4.520	0.302		0.738	
Total	17	6.121				
Bacteria						
Coast	2	1.737	0.867	3.229	0.301	**0.001**
Residuals	15	4.035	0.269		0.699	
Total	17	5.773				

B

	Df	SumsOfSqs	MeanSqs	Pseudo F	R^2	P value
COI						
Coast	2	1.632	0.816	2.083	0.217	**0.001**
Residuals	15	5.879	0.392		0.784	
Total	17	7.510				
18S						
Coast	2	1.192	0.596	1.358	0.153	**0.001**
Residuals	15	6.580	0.439		0.845	
Total	17	7.772				
16S						
Coast	2	1.670	0.835	2.712	0.267	**0.001**
Residuals	15	4.618	0.308		0.734	
Total	17	6.287				

3.8.6 Supplementary Table 4

PERMANOVA model pairwise comparison outputs based on a Jaccard index of amplicon sequence variants from eDNA metabarcoding of sites across South Africa. Models are presented for ***A*** *three taxonomic groups: metazoa, protists and bacteria.* ***B*** *three genes: COI – cytochrome c oxidase subunit I; 18S – nuclear small subunit ribosomal DNA; 16S – prokaryotic small subunit ribosomal DNA. Significant p-values at α=0.05 level are indicated in* ***bold****.*

A

	Df	SumsOfSqs	F.Model	R2	p.value	p.adjusted
Metazoa						
East vs South	1	0.671	1.853	0.156	0.002	**0.005**
East vs West	1	0.859	2.222	0.182	0.005	**0.005**
South vs West	1	0.737	1.861	0.157	0.005	**0.005**
Protists						
East vs South	1	0.800	2.843	0.221	0.004	**0.004**
East vs West	1	0.885	2.887	0.224	0.004	**0.004**
South vs West	1	0.717	2.266	0.185	0.003	**0.004**
Bacteria						
East vs South	1	0.738	3.135	0.239	0.001	**0.003**
East vs West	1	1.093	3.993	0.285	0.003	**0.003**
South vs West	1	0.775	2.602	0.206	0.003	**0.003**

B

	Df	SumsOfSqs	F.Model	R2	p.value	p.adjusted
COI						
East vs South	1	0.705	1.723	0.147	0.002	**0.006**
East vs West	1	0.899	2.232	0.182	0.003	**0.009**
South vs West	1	0.844	2.322	0.188	0.001	**0.003**
18S						
East vs South	1	0.612	1.411	0.124	0.016	**0.048**
East vs West	1	0.615	1.374	0.121	0.002	**0.006**
South vs West	1	0.560	1.290	0.114	0.006	**0.018**
16S						
East vs South	1	0.732	2.180	0.179	0.007	**0.021**
East vs West	1	1.081	3.411	0.254	0.004	**0.012**
South vs West	1	0.691	2.553	0.203	0.002	**0.006**

3.8.7 Supplementary Table 5

Analysis of variance model output for tests of multivariate homogeneity of group dispersion (betadisp R function) based on a Jaccard index of amplicon sequence variants from eDNA metabarcoding of sites across South Africa. Models are presented for ***A*** *three taxonomic groups: metazoa, protists and bacteria.* ***B*** *three genes: COI – cytochrome c oxidase subunit I; 18S – nuclear small subunit ribosomal DNA; 16S – prokaryotic small subunit ribosomal DNA. Significant p-values at $\alpha=0.05$ level are indicated in* ***bold.***

A

	Df	SumsOfSqs	MeanSq	F.Value	P
Metazoa					
Groups	2	0.008469	0.0042347	1.9485	0.177
Residuals	15	0.032599	0.0021733		
Protists					
Groups	2	0.012200	0.0061001	1.4672	0.262
Residuals	15	0.062364	0.0041576		
Bacteria					
Groups	2	0.038883	0.0194414	4.0874	**0.038**
Residuals	15	0.071346	0.0047564		

TukeyPostHoc (Bacteria)	Comparison	diff	lwr	upr	P.adj
	South vs West	0.04778014	-0.0556455	0.1512058	0.471
	East vs West	0.11338036	0.00995475	0.216806	**0.031**
	East vs South	0.06560021	-0.0378254	0.1690258	0.257

B

	Df	SumsOfSqs	MeanSq	F.Value	P
COI					
Groups	2	0.016455	0.0082276	2.3376	0.1307
Residuals	15	0.052731	0.0035154		
18S					
Groups	2	0.0014773	0.00073863	0.5964	0.5633
Residuals	15	0.0185756	0.00123837		
16S					
Groups	2	0.037211	0.0186056	3.9383	**0.0422**
Residuals	15	0.070864	0.0047243		

3.8.8 Supplementary Table 6

Mantel test summary output for both partial Mantel and corrected Mantel tests. Parameters are as follows, SST – mean sea surface temperature (°C); SSS – mean sea surface salinity (parts per thousand); Chl a – chlorophyll a concentration (mg m^{-3}); impact – human marine impact score (unitless measurement, see details in main text). Models are presented for ***A*** *three taxonomic groups: metazoa, protists and bacteria.* ***B*** *three genes three genes: COI – cytochrome c oxidase subunit I; 18S – nuclear small subunit ribosomal DNA; 16S – prokaryotic small subunit ribosomal DNA. Significant p-values at $\alpha=0.05$ level are indicated in* ***bold.***

A

	SST		SSS		ChlA		Impact	
	R statistic	p value	R statistic	p value	R statistic	p value	R statistic	p value
Metazoans partial	0.264	**0.013**	-0.148	0.881	0.066	0.272	0.432	**0.001**
Metazoans corrected	0.175	**0.015**	-0.031	0.65	0.071	0.235	0.308	**0.002**
Protists partial	0.135	0.073	0.026	0.388	0.214	**0.012**	0.317	**0.001**
Protists corrected	0.129	0.066	0.04	0.352	0.165	**0.04**	0.249	**0.01**
Bacteria partial	0.304	**0.008**	-0.133	0.843	0.021	0.403	0.492	**0.001**
Bacteria corrected	0.207	**0.007**	-0.001	0.536	0.062	0.294	0.326	**0.002**

B

	SST		SSS		ChlA		Impact	
	R statistic	**p value**	**R statistic**	**p value**	**R statistic**	**p value**	**R statistic**	**p value**
Partial Mantel COI	0.323	**0.002**	0.002	0.484	0.077	0.195	0.357	**0.003**
Corrected Mantel COI	0.157	**0.022**	0.005	0.531	0.057	0.277	0.234	**0.007**
Partial Mantel 18S	0.143	**0.059**	−0.024	0.602	0.139	0.076	0.394	**0.001**
Corrected Mantel 18S	0.162	**0.03**	0.041	0.355	0.134	0.087	0.297	**0.003**
Partial Mantel 16S	0.304	**0.008**	−0.133	0.842	0.021	0.407	0.492	**0.001**
Corrected Mantel 16S	0.207	**0.007**	0	0.531	0.062	0.302	0.326	**0.002**

3.8.9 Supplementary Table 7

Distance based redundancy analysis model outputs for models with Jaccard dissimilarities from environmental DNA metabarcoding data from South Africa as the response variable. Explanatory variables are as follows, SST – mean sea surface temperature (°C); Chl a – chlorophyll a concentration (mg m^{-3}); impact – human marine impact score (unitless measurement, see details in main text). Models are presented for datasets composed of metazoa, protists and bacteria in the COI, 18S and 16S datasets respectively. Additional model outputs are shown for the metazoan portion of the 18S dataset and the protist portion of the COI dataset. Significant p-values at α=0.05 level are indicated in ***bold****.*

	Df	SumOfSqs.	F Value	P Value
Metazoa				
SST	1	0.8328	2.1424	**1.00E-04**
Impact	1	0.5165	1.3286	**0.0277**
Residual	15	5.8308		
		Adj.R2	0.086	
Protists				
Chl a	1	0.7888	2.4641	**1.00E-04**
Impact	1	0.5301	1.656	**0.009299**
Residual	15	4.802		
		Adj.R2	0.111	
Bacteria				
	Df	SumOfSqs	F	Pr(>F)
SST	1	1.0178	3.655	**1.00E-04**
Impact	1	0.5194	1.8652	**0.009299**
Residual	15	4.1769		
		Adj.R2	0.18	
Metazoa (18S)				
	Df	SumOfSqs	F	Pr(>F)
SST	1	0.823	2.2557	**1.00E-04**
Impact	1	0.6059	1.6607	**0.0021**
Residual	15	5.4728		
		Adj.R2	0.111	

Protists (COI)

	Df	SumOfSqs	F	Pr(>F)
SST	1	1.2049	3.9839	**1.00E-04**
Impact	1	0.5637	1.8638	**0.0171**
Residual	15	4.5366		
		Adj.R2	0.192	

3.8.10 Supplementary Table 8

Generalised additive model with a restricted maximum likelihood (REML) 2D smoother summary, each row represents an individually fitted model. Parameters are as follows, SST – mean sea surface temperature (°C); SSS – mean sea surface salinity (parts per thousand); Chl a – chlorophyll a concentration (mg m^{-3}); impact – human marine impact score (unitless measurement, see details in main text). Models are presented for ***A*** *three taxonomic groups: metazoa, protists and bacteria.* ***B*** *three genes: COI – cytochrome c oxidase subunit I; 18S – nuclear small subunit ribosomal DNA; 16S – prokaryotic small subunit ribosomal DNA. Significant p-values at $\alpha=0.05$ level are indicated in* ***bold.***

A

	Parametric Coef.				Smoother			
	Estimate	Std.Error	T Value	P Value	Df	F	P Value	Adj.R2
Metazoa								
Impact	3.2348	0.1593	20.31	2.08E-10	9	3.563	**0.00297**	0.654
SST	18.56054	0.09681	191.7	<2e-16	9	100.2	**<2e-16**	0.982
Protists								
impact	3.2348	0.1779	18.18	5.8E-10	9	2.481	**0.0126**	0.568
chl a	1.8003	0.05603	32.13	1.89E-14	9	16.15	2.63E-13	0.895
Bacteria								
impact	3.2348	0.1482	21.83	2.45E-10	9	4.409	**0.00207**	0.700
SST	18.56054	0.09753	190.3	<2e-16	9	98.73	**<2e-16**	0.981

B

	Parametric Coef.				Smoother			
	Estimate	Std.Error	T Value	P Value	Df	F	P Value	Adj.R2
COI								
SSS	35.225667	0.006731	5233	<2e-16	9	27.48	**3.79E-13**	0.936
impact	3.2348	0.1753	18.45	1.51E-10	9	2.608	**0.0044**	0.58
chl a	1.8003	0.05529	32.56	3.04E-13	9	16.63	**3.47E-12**	0.898
SST	18.56054	0.09966	186.2	<2e-16	9	94.47	**<2e-16**	0.98
18S								
SSS	35.225667	0.008706	4046	<2e-16	9	15.67	**5.68E-09**	0.892
impact	3.2348	0.1477	21.9	9.75E-11	9	4.45	**0.00117**	0.702
chl a	1.8003	0.04337	41.51	3.86E-16	9	28.21	**<2e-16**	0.937
SST	18.5605	0.1531	121.2	<2e-16	9	38.93	**<2e-16**	0.954
16S								
SSS	35.225667	0.007869	4476	<2e-16	9	19.6	**1.00E-10**	0.912
impact	3.2348	0.1545	20.94	2.93E-10	9	3.907	**0.00311**	0.674
chl a	1.8003	0.07156	25.16	3.76E-12	9	9.17	**1.47E-07**	0.829
SST	18.56054	0.09916	187.2	<2e-16	9	95.44	**<2e-16**	0.981

3.8.11 Supplementary Table 9

Details of the sampling sites including site code, geographic locations [West (W), South (S) and East (E)], sampling date (year 2017) and type of site (natural or artificial). Order of appearance of sites progresses along the sampled coast, from the northwest coast to the northeast coast of South Africa.

Site Name	*Site Code*	*Latitude (S)*	*Longitude (E)*	*Sampling Date*	*Geographic Coastline*	*Site Classification*
Yachtport SA, Saldanha	SY	-33.026701	17.960758	14-Oct	W	Artificial
Saldanha Natural Site	SN	-33.071672	18.038446	15-Oct	W	Natural
Club Mykonos Marina	SM	-33.046315	18.04028	15-Oct	W	Artificial
V&A Waterfront Marina	TB	-33.909055	18.419893	18-Oct	W	Artificial
Hout Bay Yacht Club	HB	-34.049767	18.348042	19-Oct	W	Artificial
Hout Bay Natural Site	HN	-34.049016	18.360954	19-Oct	W	Natural
Mossel Bay Marina	MB	-34.178386	22.144918	22-Oct	S	Artificial
Mossel Bay Natural Site	MN	-34.299278	21.944026	22-Oct	S	Natural
Knysna	KN	-34.041086	23.043592	24-Oct	S	Artificial
Knysna Natural Site	NK	-34.05877	23.034025	24-Oct	S	Natural
Cape St Francis	CN	-34.212257	24.837625	26-Oct	S	Natural
Port Elizabeth Marina	PE	-33.96692	25.634461	27-Oct	S	Artificial
Bushmans River (Kenton Marina)	BR	-33.679558	26.655841	30-Oct	E	Artificial
Port Alfred Marina	PA	-33.593546	26.892084	31-Oct	E	Artificial
East London Marina	EL	-33.024163	27.896269	02-Nov	E	Artificial
Durban Marina	DU	-29.862663	31.021903	08-Nov	E	Artificial
Richards Bay Marina	RB	-28.793908	32.079184	10-Nov	E	Artificial
Richards Bay Natural	RN	-28.827437	32.035825	11-Nov	E	Natural

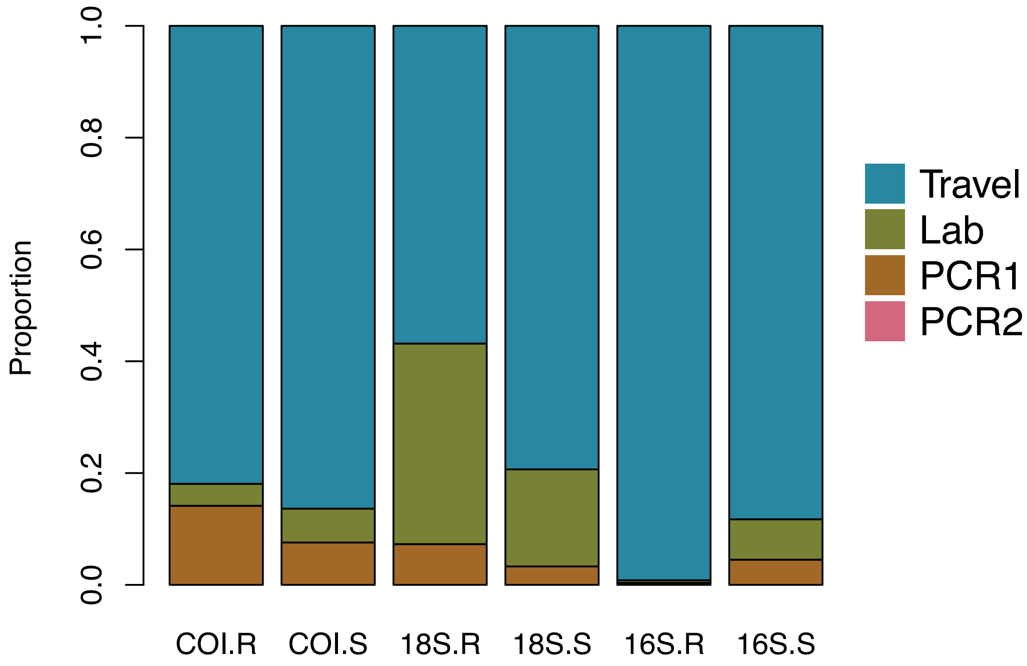
3.8.12 Supplementary Note 1

Information concerning control samples used during field sampling and in the laboratory during the construction of metabarcoding libraries for high-throughput sequencing of eDNA extracted from water samples collected along the coast of South Africa.

Each water sample was taken using a peristaltic pump, associated tubing was washed with 1 litre of seawater from the site before conducting the field sampling. During fieldwork all reused field equipment (hosing and plasticware) was washed thoroughly between sites with 5% bleach solution and rinsed with tap water.

Negative controls were implemented during each DNA extraction and during each PCR step of library construction. Across all controls for COI a total of 6.6 ± 7.6 (s.d.) ASVs were found in each control sample with a summed average of 117.3 ± 155.1 reads per sample. Across the 18S controls 12.1 ± 9.5 (s.d.) ASVs had a summed average of 579.1 ± 517.5 reads per sample. The 16S data showed the largest number of average ASVs (61.7 ± 63.1) per sample and also more summed reads on average (16,079.9 ± 31,915.2). A single ASV, assigned to the genus *Bradyrhizobium*, contributed the bulk of these reads (mean of 25,444.4 per control sample) and was found only in control samples used to assess decontamination in field. This genus of bacteria are Gram-negative soil bacteria, some of which have roles in N_2 fixation(Stacey et al. 1995) and therefore likely represent contamination from the water used for cleaning apparatus between sites. However, this genus of bacteria has also been shown to contaminate ultra-pure water and laboratory reagents(Kulakov et al. 2002, Salter et al. 2014). After bioinformatically removing this ASV the 16S control samples contained a mean of 56.1 ± 61.9 ASVs per sample and a mean of 3,054.9 ± 3,398.8 summed reads per sample.

The proportion of contamination assigned to travel, lab and PCR controls are shown below in Supplementary Note 1 Figure 1. The largest proportion of contamination was from the travel controls, indicating that the cleaning process between sites was the largest source of contamination. The level of contamination shown here is typical for eDNA studies(Holman et al. 2019, Jeunen et al. 2019, Blackman et al. 2020) and all contamination was bioinformatically filtered as detailed in the main manuscript. Therefore, we are confident that contamination did not affect the inference presented in the main manuscript.

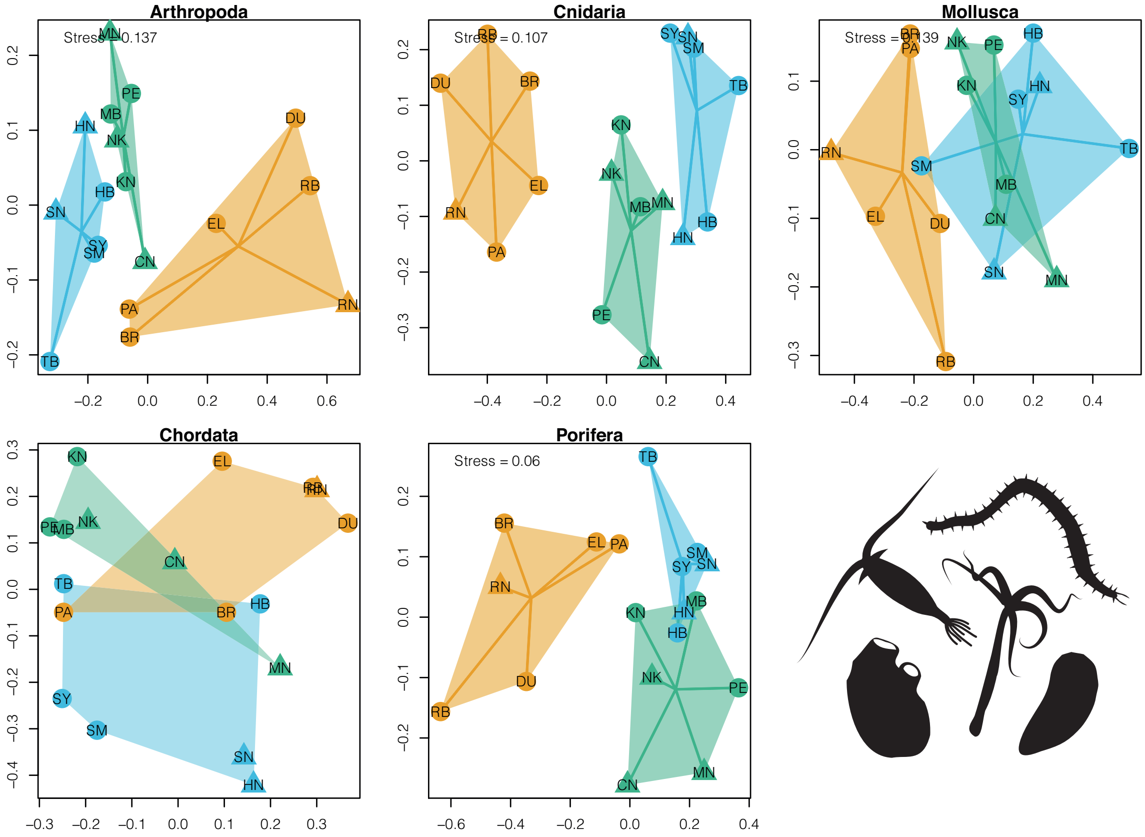


Supplementary Note 1 Figure 1 *Proportion of contamination attributed to different sections of the workflow. Each column shows the proportion of either reads (COI.R, 18.R, 16.R) or ASVs (COI.S, 18S.S, 16S.S) contributed by each of different types of control sample for all three sequenced markers (COI, 18S, 16S). Travel controls include sealed controls and tap water controls used to evaluate cleaning of equipment between sampling. Lab controls include sealed filters, DNA extraction kits and inhibition clean up kits. PCR controls are no template controls added before running the PCR.*

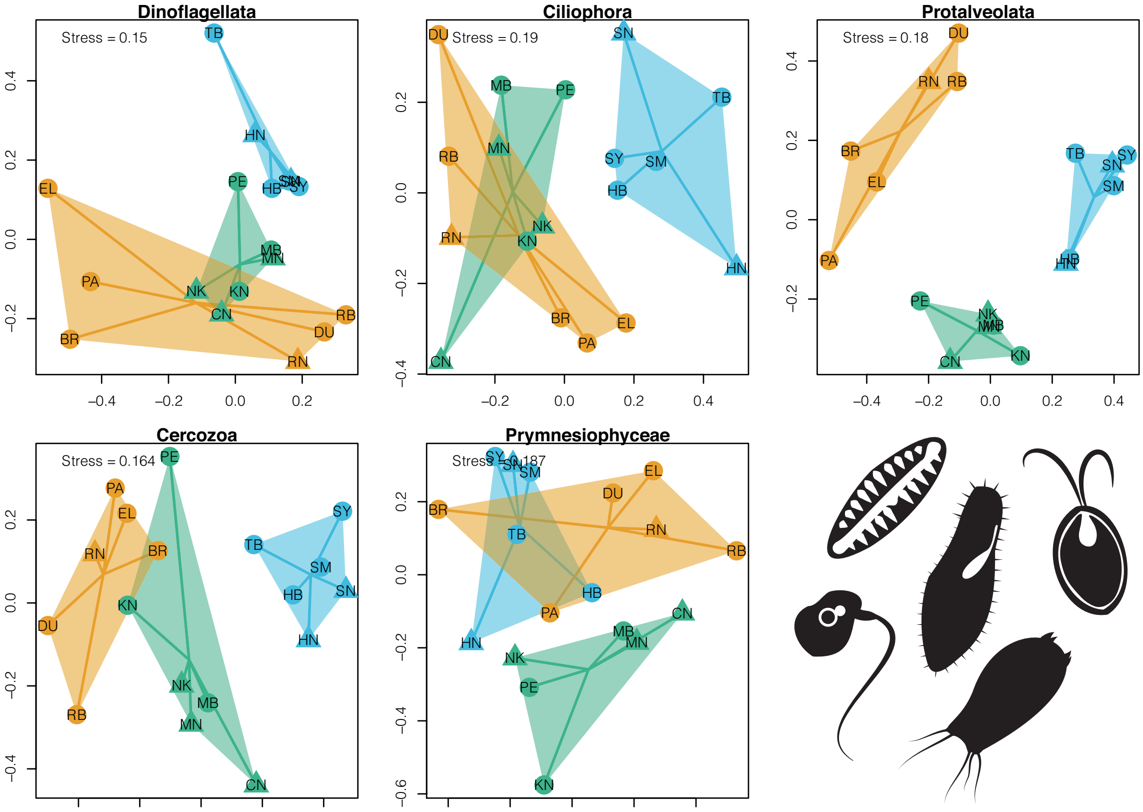
3.8.13 Supplementary Note 2

Phyla-level analysis of biogeographic patterns

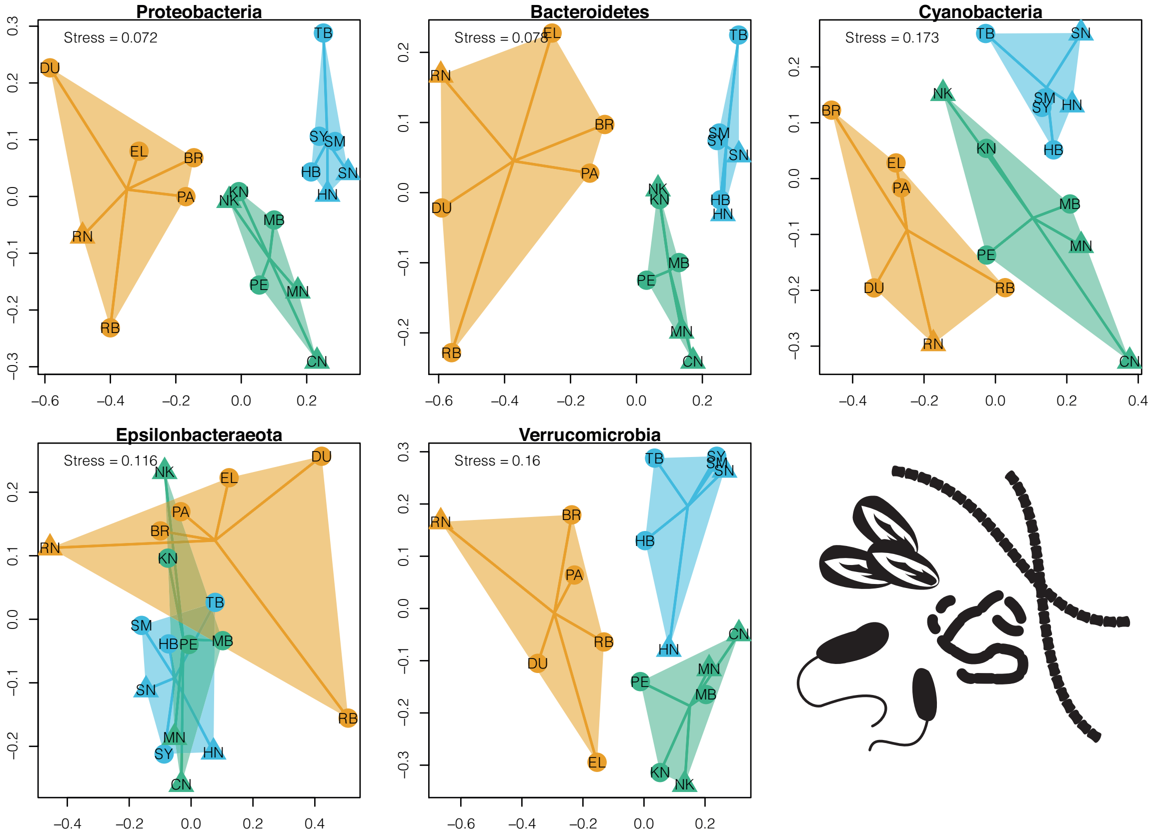
In order to evaluate the beta diversity patterns at lower taxonomic levels than kingdom each dataset was subset to include only ASVs allocated to the top five most ASV rich phyla per taxonomic dataset (metazoans, protists, bacteria). A non-metric multidimensional ordination, evaluation of multivariate heterogeneity and PERMANOVA were performed on each of these sub-datasets as detailed in the main manuscript methods for the complete datasets.



Supplementary Note 2 Figure 1 *Observed patterns of β-diversity from environmental DNA metabarcoding (COI metazoans) based on Jaccard dissimilarities between amplicon sequence variants along the coast of South Africa. Plots shows non-metric multidimensional scaling ordinations for each phylum. Coloured hulls show the spread of the data and lines indicate the spread around the centroid grouped by coast with the east, south and west coasts denoted by orange, green and blue respectively. Site name abbreviations as in Supplementary Table 9. The ordination for Chordata is the first two axes of a principle coordinate analysis due to non-convergence of an optimal nMDS solution representing the Chordata data.*



Supplementary Note 2 Figure 2 *Observed patterns of β-diversity from environmental DNA metabarcoding (18S protists) based on Jaccard dissimilarities between amplicon sequence variants along the coast of South Africa. Plots shows non-metric multidimensional scaling ordinations for each phylum. Coloured hulls show the spread of the data and lines indicate the spread around the centroid grouped by coast with the east, south and west coasts denoted by orange, green and blue respectively. Site name abbreviations as in Supplementary Table 9.*



Supplementary Note 2 Figure 3 *Observed patterns of β-diversity from environmental DNA metabarcoding (bacteria) based on Jaccard dissimilarities between amplicon sequence variants along the coast of South Africa. Plots shows non-metric multidimensional scaling ordinations for each phylum. Coloured hulls show the spread of the data and lines indicate the spread around the centroid grouped by coast with the east, south and west coasts denoted by orange, green and blue respectively. Site name abbreviations as in Supplementary Table 9.*

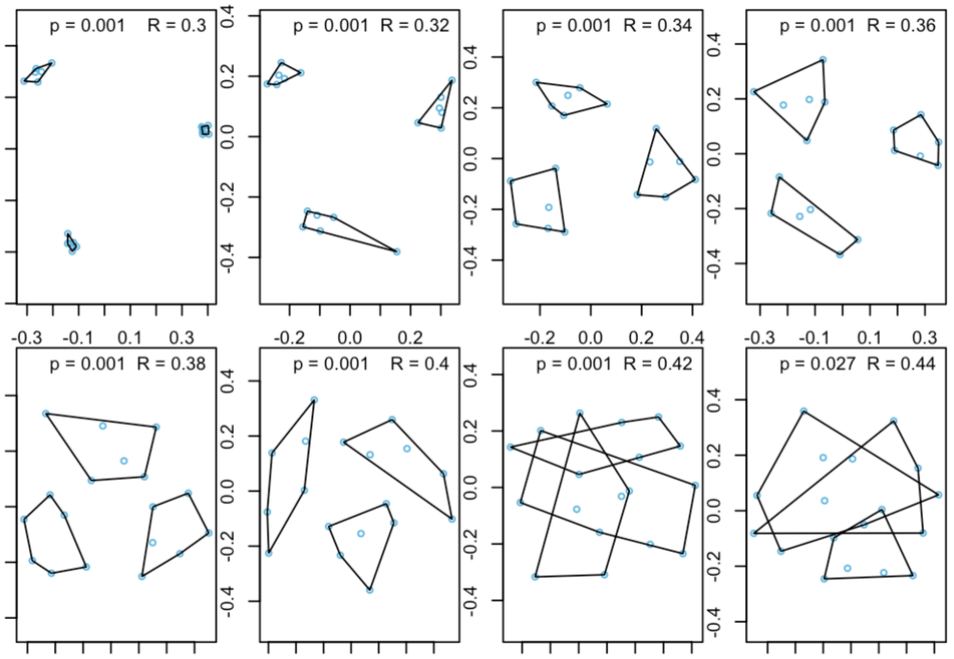
Supplementary Note 2 Table 1 *Phyla level dataset outputs showing number of ASVs per phyla (nASVs), the outputs (F statistic and p value) from an analysis of variance on multivariate distribution between ecoregions for each dataset (permdisp from vegan in R). F statistic and P value from a PERMANOVA model based on Jaccard dissimilarities testing a difference in between ecoregions.*
Significant p-values at $\alpha=0.05$ *level are indicated in* ***bold.***

			Multivariate Dispersion		**PERMANOVA**	
Phyla	**Dataset**	**nASVs**	**Fstat**	**p value**	**Fstat**	**p value**
Arthropoda	Metazoa	354	1.742	0.209	1.816	**0.001**
Cnidaria	Metazoa	258	0.774	0.479	2.468	**0.001**
Mollusca	Metazoa	106	0.377	0.692	2.07	**0.001**
Chordata	Metazoa	92	0.241	0.789	1.452	**0.001**
Porifera	Metazoa	68	2.771	0.095	2.261	**0.001**
Dinoflagellata	Protist	257	2.637	0.104	3.086	**0.001**
Ciliophora	Protist	389	0.437	0.654	2.024	**0.001**
Protalveolata	Protist	281	1.958	0.176	2.802	**0.001**
Cercozoa	Protist	218	1	0.391	2.446	**0.001**
Prymnesiophyceae	Protist	70	0.178	0.839	2.598	**0.001**
Proteobacteria	Bacteria	1340	2.489	0.117	3.208	**0.001**
Bacteroidetes	Bacteria	673	7.38	**0.006**	4.084	**0.001**
Cyanobacteria	Bacteria	390	2.776	0.094	2.826	**0.001**
Epsilonbacteraeota	Bacteria	83	2.158	0.15	1.47	**0.029**
Verrucomicrobia	Bacteria	84	0.45	0.646	2.748	**0.001**

3.8.14 Supplementary Note 3

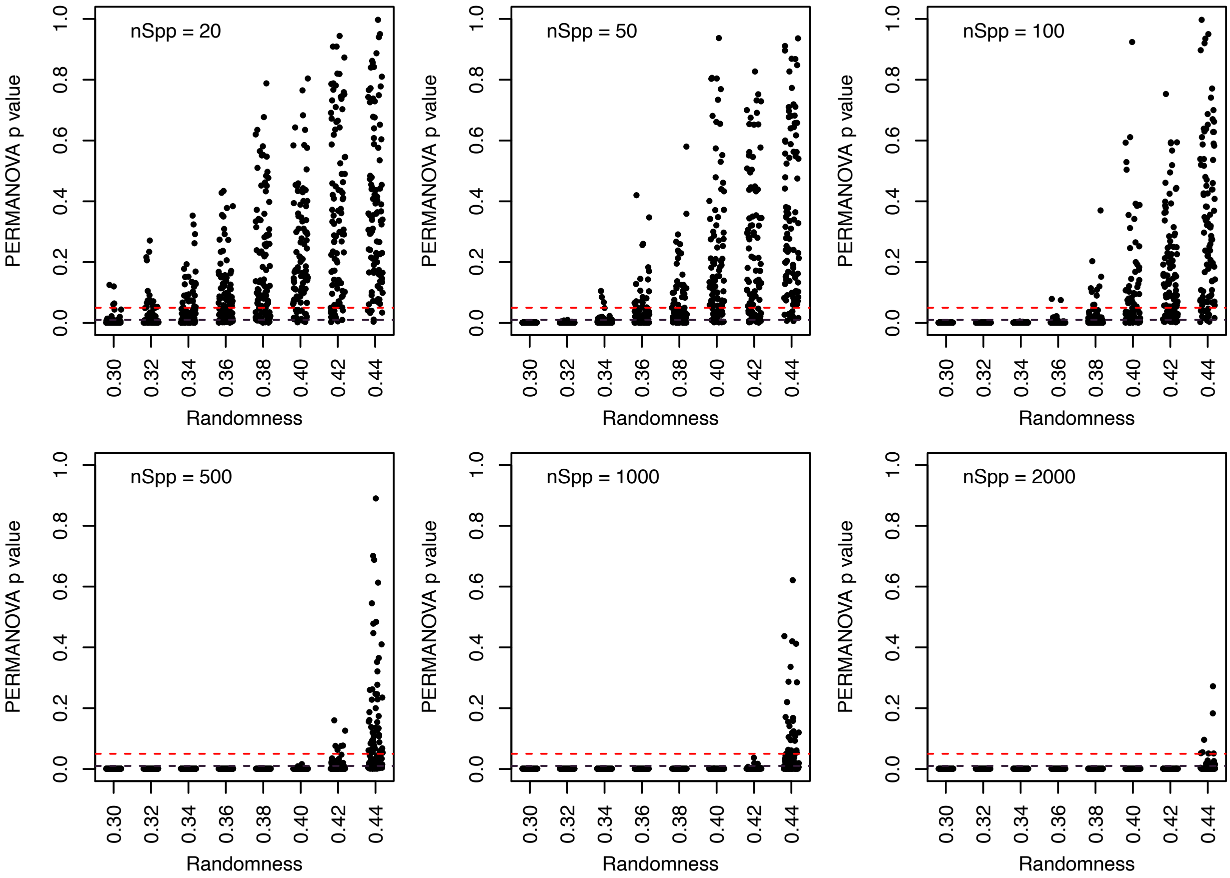
Simulations of power for detecting differences in eco-regions using PERMANOVAs.

We tested how many species observations were required to detect a significant difference between the multivariate spread of the three ecoregions (see Methods, main text). This was important as the number of ASVs subsets at phyla level can be low (see Supplementary Note 2 Table 1). Therefore, we conducted a community simulation using data simulated using random distributions in R (v4.0.2) as follows. A randomly-generated dataset containing 1,000 simulated species was created. Three possible types of spatial distribution were possible for each species: 1. Panmixia (i.e. the species is present in all sites), 2. Random distribution (randomly distributed across sites), and 3. ecoregion-determined species (species found only in one of the three ecoregions). Species occurrence was coded as 1 (present) or 0 (absent) across 18 simulated sites in three different ecoregions with a fixed proportion of species for each of the above types (7:2:1 ecoregion:random:panmixia) spread across the ecoregions (7:2:1 ecoregion:random:panmixia). These observations were then subject to incidence values changes (1 changed to 0 or 0 changed to 1) at varying proportions across the entire dataset (30% to 44% of data), these proportions produced ecoregion separation on nMDS plots similar to empirical ecological datasets during initial trials. The output of this simulation was visualised using an nMDS ordination of Jaccard dissimilarities (see Supplementary Note 3 Figure 1).



Supplementary Note 3 Figure 1 *Non-metric multidimensional scaling ordinations based on Jaccard dissimilarities of simulated community data. Simulated communities had 18 sites equally spaced among three ecoregions and 1000 species were simulated. A proportion of species incidence records were randomised from 30% (R=0.3) to 44% (R=0.44). Sites are shown in blue and a convex hull is shown for each ecoregion. The p value reported in each simulation corresponds to the PERMANOVA result testing for significant differences in the multivariate spread of ecoregions.*

A PERMANOVA was used to assess the significance of differences in multivariate location between the three ecoregions. As an ecoregion difference was simulated in the community data a p-value greater than 0.05 indicates a false negative result due to stochasticity. This simulation was replicated 100 times for 20, 50, 100, 500, 1000 and 2000 species (Supplementary Note 3 Figure 2).



Supplementary Note 3 Figure 2 *Plots showing the significance value (p value) for a PERMANOVA testing multivariate location between three ecoregions in simulated community data. Simulations were conducted for eight levels of randomness values: indicating the proportion of observations changes to an alternative state (present to absent or vice versa). Each plot shows 100 simulations for each level of randomness, with different plots showing effect for an increasing number of simulated species observations (nSpp). The red dashed lines indicate a PERMANOVA p-value of 0.05 and the dark purple dashed line 0.01.*

The ordinations between 0.32 and 0.38 showed the highest similarity to those presented in the main manuscript (Fig. 2). The proportion of false negative simulations was between 0.15-0.71 for 20 species and 0-0 for 2000 species (see Supplementary Note 3 Table 1). Between 50 and 100 species the false negative proportion decreased from an average of 0.13 to 0.03, representing a shift from 87% to 97% of simulations finding a significant effect of ecoregions. Therefore, a reasonably high (0.95) chance of seeing a true effect lies between 50 and 100 species.

Supplementary Note 3 Table 1 *Proportion of false negative results (p >0.05) from a PERMANOVA significance test of ecoregion across 100 simulated datasets. Each dataset was simulated for a given randomness value (0.3-0.44 – shown on left) and number of species (20-2000).*

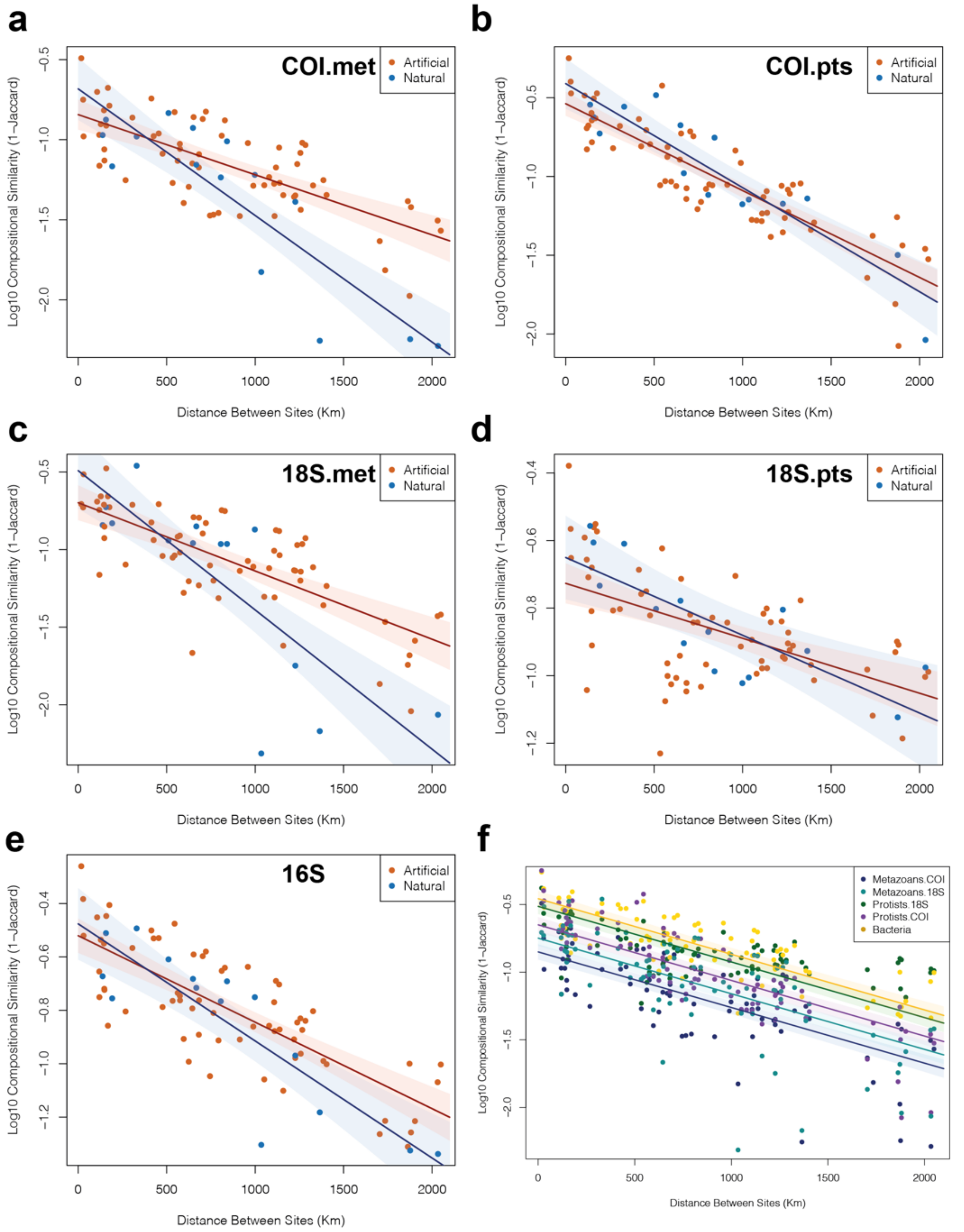
	Number of Species					
Randomness	**20**	**50**	**100**	**500**	**1000**	**2000**
0.3	0.04	0	0	0	0	0
0.32	0.15	0	0	0	0	0
0.34	0.28	0.03	0	0	0	0
0.36	0.53	0.21	0.02	0	0	0
0.38	0.71	0.28	0.1	0	0	0
0.4	0.8	0.72	0.39	0	0	0
0.42	0.87	0.69	0.71	0.06	0	0
0.44	0.92	0.9	0.86	0.56	0.26	0.07

3.8.15 Supplementary Note 4

Distance-decay relationship analyses

Analysis by taxonomic group

Distance-decay slopes for all observations showed an exponential decrease in compositional similarity as the distance between sites increased (Supplementary Note 4 Figure 1, see below). Regression models of log10 transformed compositional similarity indicated that this slope was statistically significant in all cases ($p < 0.001$ for all markers & taxonomic groups, full model output below). In the metazoan datasets (both COI & 18S) the model showed a significant difference in the slope between artificial and natural sites (COI - $p = 0.0002$; 18S - $p = 0.0019$, full model outputs below). No statistically significant differences were found between site types in the protist or bacteria datasets.



Supplementary Note 4 Figure 1 *Plots showing distance between sites and community dissimilarity measured by environmental DNA metabarcoding across South Africa. Logarithmically (base 10) transformed compositional similarity against distance is shown for **a** COI metazoans, **b** COI protists, **c** 18S metazoans **d** 18S protists and **e** 16S data with comparisons between artificial sites coloured red and natural sites coloured blue. 95% confidence intervals from the least-squares linear regression models are shown as light shaded areas* around each regression slope. **f** All distance decay data overlayed.

Supplementary Note 4 Table 1 *Least-squares regression model for Distance-Decay relationships between different site types measured by eDNA metabarcoding for coastal sites across South Africa. Models are presented for three genes: COI – cytochrome c oxidase subunit I; 18S – nuclear small subunit ribosomal DNA; 16S – prokaryotic small subunit ribosomal DNA.*

COI metazoan				
$F_{(3,76)}$=47.73				
R2 =0.6396	Estimate	Std.Error	t value	P value
(Intercept)	-8.43E-01	4.71E-02	-17.922	**<2e-16**
distance	-3.76E-04	4.77E-05	-7.88	**1.87E-11**
site.typeNatural	1.61E-01	1.08E-01	1.488	0.140856
distance:site.typeNatural	-4.15E-04	1.06E-04	-3.905	**0.000202**

COI protists				
$F_{(3,77)}$=94.5				
R2 =0.7781				
(Intercept)	-5.38E-01	3.82E-02	-14.068	**<2e-16**
distance	-5.52E-04	3.79E-05	-14.534	**<2e-16**
site.typeNatural	1.28E-01	8.86E-02	1.441	0.154
distance:site.typeNatural	-1.09E-04	8.67E-05	-1.261	0.211

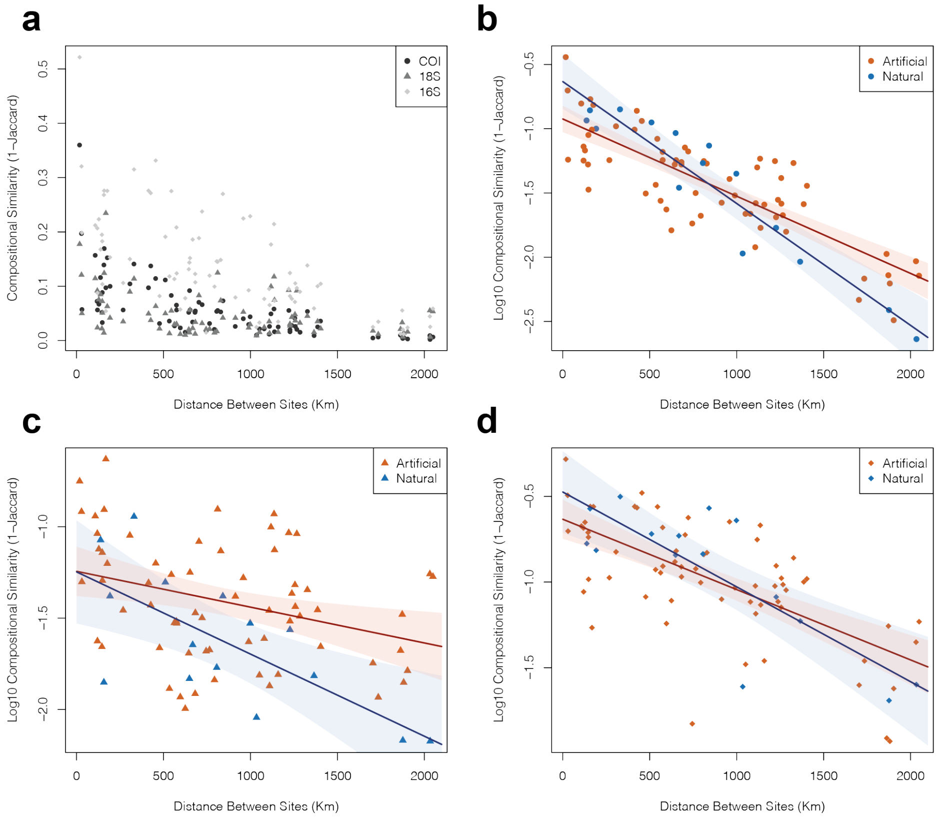
18S metazoan				
$F_{(3,76)}$=37.22				
R2 =0.579	Estimate	Std.Error	t value	P value
(Intercept)	-6.98E-01	5.63E-02	-12.394	**<2e-16**
distance	-4.40E-04	5.59E-05	-7.867	**1.98E-11**
site.typeNatural	2.07E-01	1.35E-01	1.534	0.12916
distance:site.typeNatural	-4.58E-04	1.43E-04	-3.213	**0.00193**

18S protists				
$F_{(3,77)}$=15.09				
R2 =0.3458				
(Intercept)	-7.27E-01	2.97E-02	-24.504	**<2e-16**
distance	-1.63E-04	2.94E-05	-5.528	**4.25E-07**
site.typeNatural	7.65E-02	6.87E-02	1.114	0.269
distance:site.typeNatural	-6.76E-05	6.72E-05	-1.006	0.318

16S				
$F_{(3,77)}$=25.82				
R2 =0.4821				
(Intercept)	-1.52E-01	9.04E-03	-16.826	**<2e-16**
distance	7.16E-05	8.97E-06	7.978	**1.12E-11**
site.typeNatural	1.22E-02	2.09E-02	0.583	0.562
distance:site.typeNatural	-4.68E-06	2.05E-05	-0.229	0.82

Analysis by gene region

We analysed the distance-decay slopes for all observations and these showed an exponential decrease in compositional similarity as the distance between sites increased (Supplementary Note 4 Figure 2 below). Regression models of log10 transformed compositional similarity indicated that this slope was statistically significant in all cases ($P<0.005$ for all markers, full model output below). In the COI data the model showed a significant difference in both the slope and intercept between artificial and natural sites ($F_{3,77}=77.92$, additive $p=0.018$, interactive $p=0.004$, full model below). No statistically significant difference was found between site types in the 18S or 16S data. As shown in Fig. 3b-d, there was much greater variance in the residuals for the 18S and 16S data compared to the COI data and the R^2 values from the models support these observations (COI - $R^2=0.730$, 18S - $R^2=0.224$, 16S - $R^2=0.473$).



Supplementary Note 4 Figure 2. *Plots showing distance between sites and community dissimilarity measured by environmental DNA metabarcoding across South Africa.* ***a*** *Untransformed data for all gene regions. Logarithmically (base 10) transformed compositional similarity against distance is shown for* ***b*** *COI,* ***c*** *18S and* ***d*** *16S data with comparisons between artificial sites coloured red and natural sites coloured blue. 95% confidence intervals from the least-squares linear regression models are shown as light shaded areas* around each regression slope.

***Supplementary Note 4 Table 2** Least-squares regression model for Distance-Decay relationships between different site types measured by eDNA metabarcoding for coastal sites across South Africa. Models are presented for three genes: COI – cytochrome c oxidase subunit I; 18S – nuclear small subunit ribosomal DNA; 16S – prokaryotic small subunit ribosomal DNA.*

COI				
$F_{(3,77)}$=77.92				
R2 =0.7295	Estimate	Std.Error	t value	P value
(Intercept)	-9.23E-01	5.17E-02	-17.849	**<2.00E-16**
distance	-6.02E-04	5.13E-05	-11.726	**<2.00E-16**
site.typeNatural	2.89E-01	1.20E-01	2.415	**0.01812**
distance:site.typeNatural	-3.48E-04	1.17E-04	-2.967	**0.00401**

18S				
$F_{(3,77)}$=8.679				
R2 =0.2236				
(Intercept)	-1.2434844	0.0677062	-18.366	**<2e-16**
distance	-0.0001962	0.0000672	-2.919	**0.0046**
site.typeNatural	-0.0015438	0.1568279	-0.01	0.9922
distance:site.typeNatural	-0.0002571	0.0001535	-1.675	0.0979

16S				
$F_{(3,77)}$=24.94				
R2 =0.4731				
(Intercept)	-6.33E-01	5.72E-02	-11.076	**<2.00E-16**
distance	-4.10E-04	5.68E-05	-7.219	**3.19E-10**
site.typeNatural	1.58E-01	1.33E-01	1.195	0.236
distance:site.typeNatural	-1.44E-04	1.30E-04	-1.107	0.272

3.8.16 Supplementary Note 5

Quality control method and results for detecting non-biological (chimeras, unknown primer dimers, etc.) and biological (numts, pseudogenes, non-region binding) erroneous amplicon sequence variants (ASVs) in metabarcoding of eDNA data from South African seawater.

Open reading frame trial

Each ASV in the raw (post DADA2 pipeline) and quality controlled (post LULU and read frequency filtering) COI dataset was translated at each of three reading frames under all possible translation tables. The smallest number of stop codons for each frame/table was recorded per ASV. In addition, the number of stop codons per ASV for the invertebrate mitochondrial codon table was recorded to evaluate the possibility that using the minimum value across all tables ignores true stop codons as a result of using a table with fewer stop codons (e.g. NCBI Table 14 – the alternative flatworm table). The number of stop codons per ASV and percentage of ASVs with no stop codon is shown in Supplementary Note 5 Table 1 below.

Supplementary Note 5 Table 1 *Number of stop codons found in ASVs from COI eDNA metabarcoding data. Results are shown for quality controlled (QC) and raw datasets following the invertebrate codon table (invert) and the codon table with the lowest number of stop codons (MinStop). For each dataset and codon table the percentage of ASVs with no stop codon are shown.*

	QC Dataset		Raw Dataset	
Nstop Codons	MinStop	Invert	MinStop	Invert
0	4866	4845	37242	36923
1	1	11	72	138
2	0	5	2	100
3	0	5	0	85
4	0	0	0	54
5	0	1	0	15
6	0	0	0	0
7	0	0	0	1
Percentage not stop	99.98	99.55	99.80	98.95

Overall, the proportion of sequences in the QC dataset with a stop codon was negligible, indicating that only a minute fraction of the observed sequences are non-coding.

Manual BLAST trial

A random 5% subset of ASVs (243 sequences) in the quality-controlled cytochrome c oxidase I (COI) dataset was extracted. These ASVs were compared to the NCBI nt/nr database using the megablast algorithm on the online NCBI blast tool (blast.ncbi.nlm.nih.gov, accessed online 07.04.2020). Each sequence was manually examined, the top 50 hits were visually compared, and the best match (highest E value above 75% coverage) was recorded. Any ASV with a less than 70% coverage by a subject sequence was treated as a putative chimera. The query section not matching the subject in these cases was extracted and used as a query in a second search. A true chimera would have different biological hits for each section. For ASVs where no hit was found in the above database, a blastx search (nucleotide -> protein search) was performed under the standard genetic code. Protein sequence matches were expected to be cytochrome c oxidase I (COI), hits with less than 97% positive ID (amino acid matches that are identical or have conservative substitutions) were designated a 'low' quality match.

Out of 243 sequences, 20 (8.2%) could be assigned a high-quality hit of 97% sequence identity or greater. 216 (88.8%) could be assigned a poor-quality match indicating a species or taxa not represented in the sequence database. 2 (0.8%) sequences had no nucleotide match but a high-quality protein match indicating a functioning COI gene. Five sequences (2.0%) had no nucleotide match and no high-quality protein match, however all five sequences received protein matches to a COI subject indicating biological rather than technical origin. No sequences had a partial (<70%) match indicative of a chimeric sequence. Overall, this evidence suggests that the bioinformatic parameters and filtering steps used to produce the quality-controlled dataset resulted in minimal retained sequences that correspond with technical errors and non-target biological regions.

3.8.17 Supplementary References

Blackman, R. C., K. K. S. Ling, L. R. Harper, P. Shum, B. Hanfling, and L. Lawson-Handley. 2020. Targeted and passive environmental DNA approaches outperform established methods for detection of quagga mussels, *Dreissena rostriformis bugensis* in flowing water. Ecology and Evolution **10**:13248-13259.

Holman, L. E., M. de Bruyn, S. Creer, G. Carvalho, J. Robidart, and M. Rius. 2019. Detection of introduced and resident marine species using environmental DNA metabarcoding of sediment and water. Scientific Reports **9**:11559.

Jeunen, G. J., M. Knapp, H. G. Spencer, M. D. Lamare, H. R. Taylor, M. Stat, M. Bunce, and N. J. Gemmell. 2019. Environmental DNA (eDNA) metabarcoding reveals strong discrimination among diverse marine habitats connected by water movement. Molecular Ecology Resources **19**:426-438.

Kulakov, L. A., M. B. McAlister, K. L. Ogden, M. J. Larkin, and J. F. Hanlon. 2002. Analysis of Bacteria Contaminating Ultrapure Water in Industrial Systems. Applied and environmental microbiology **68**:1548.

Salter, S. J., M. J. Cox, E. M. Turek, S. T. Calus, W. O. Cookson, M. F. Moffatt, P. Turner, J. Parkhill, N. Loman, and A. W. Walker. 2014. Reagent contamination can critically impact sequence-based microbiome analyses. BMC Biology **12**.

Stacey, G., J. Sanjuan, S. Luka, T. Dockendorff, and R. W. Carlson. 1995. Signal exchange in the *Bradyrhizobium*-soybean symbiosis. Soil Biology and Biochemistry **27**:473-483.

Chapter 4 Managing human mediated range shifts: understanding spatial temporal and genetic variation in marine non-native species

4.1 Citation

Here I present a submitted manuscript, a pre-print has been archived as below.

Holman LE, Parker -Nance S, de Bruyn M, Creer S, Carvalho G, Rius M (2021) Managing human mediated range shifts: understanding spatial temporal and genetic variation in marine non-native species. *BioRxiv* 2021.07.06.451303. DOI:10.1101/2021.07.06.451303.

4.2 Abstract

The use of molecular methods to manage natural resources is increasingly common. However, DNA-based methods are seldom used to understand the spatial and temporal dynamics of species' range shifts. This is important when managing range-shifting species such as non-native species (NNS), which can have negative impacts on biotic communities. Here we investigated the range-shifting NNS *Ciona robusta, Clavelina lepadiformis, Microcosmus squamiger* and *Styela plicata* using a combined methodological approach. We first conducted non-molecular biodiversity surveys for these NSS along the South African coastline, and compared the results with historical surveys. We detected no consistent change in range size across species, with some displaying range stability and others showing range shifts. We then sequenced a section of cytochrome c oxidase subunit I (COI) from tissue samples and found genetic differences along the coastline but no change over recent times. Finally, we found that environmental DNA metabarcoding data showed broad congruence with both the non-molecular biodiversity and the COI datasets, but failed to capture complete incidence of all NSS. Overall, we demonstrated how a combined methodological approach can effectively detect spatial and temporal variation in genetic composition and range size, which is key for managing biodiversity changes of both threatened and NSS.

4.3 Introduction

Biodiversity is undergoing a global redistribution as a result of human influence, with species increasingly found in environments outside their previously reported geographic range (Pecl et al. 2017). Contemporary climate change is causing species to shift their ranges to accommodate novel environmental conditions (Sunday et al. 2012, Sunday et al. 2015), and human-mediated species introductions dramatically increase the range of non-native species (NNS) (Bax et al. 2003, Molnar et al. 2008, Seebens et al. 2021). This exposes species to abiotic conditions and biotic interactions that are different to those experienced in native habitats. Such changes in distribution can result in a dramatic increase or decrease in population size, or may have a limited detectable immediate effect (Pecl et al. 2017, Dornelas et al. 2019). Understanding these responses is important to answer fundamental ecological and evolutionary questions about changing biotic communities, but also for natural resource managers when predicting changes in ecosystem services and natural capital (Vilà et al. 2010, Pecl et al. 2017).

Global biodiversity loss has consistently been shown to reduce ecosystem function, and in turn affects the provision of ecosystem services (Cardinale et al. 2012, Gamfeldt et al. 2015). A key driver of biodiversity loss is the introduction of NNS (Mazor et al. 2018), which also imposes a substantial global economic cost (Diagne et al. 2021) and has a dramatic impact on public health (Mazza et al. 2014, Schindler et al. 2015). In the marine environment the majority of NNS introductions are associated with transoceanic shipping (Molnar et al. 2008, Katsanevakis et al. 2013, Williams et al. 2013) and therefore, major ports and harbours are hotspots for NNS. Once a species is introduced to these sites, secondary spread can be facilitated by smaller recreational vessels, marinas and marine infrastructure surrounding these major harbours (Glasby et al. 2007, Airoldi et al. 2015). Considering the increasing number of yearly NNS introductions (Seebens et al. 2017, Seebens et al. 2021), improving our understanding of how range shifts of NNS occur through time and space is critical in the design of effective management and mitigation responses.

Natural resource managers have finite budgets and limited information when making decisions simultaneously on a number of NNS with variable or unknown impact (Roy et al. 2019, Booy et al. 2020). For each NNS, managers can attempt to eradicate a population, make efforts to avoid any further expansion into new areas, or acknowledge that control is not possible and work on mitigation strategies (Thresher and Kuris 2004, Clout and Williams 2009). These limited options are compounded by the vast costs associated with control or eradication, and even when control methods may be possible, they might be politically or publicly unacceptable (Liordos et al. 2017, Shackleton et al. 2019). Furthermore, control measures can be unsuccessful because of incomplete eradication of the target species or ongoing species reintroductions (Pluess et al. 2012, Booy et al. 2020, Simberloff 2020). Consequently, managers frequently take no action to control NNS or act only when evidence for both presence and substantial impact has been gathered (Giakoumi et al. 2019). It is therefore beneficial to develop tools that provide researchers and managers with information to facilitate decision-making. Genetic tools can complement existing methods for assessing NNS range shifts by providing information that would be unfeasible or impossible to produce otherwise (Darling et al. 2017).

Even when NNS can be unambiguously identified, it can be difficult to determine when and where they were first introduced into a region, (for example see Hudson et al. (2021)). Since eradication or control efforts are improved by early detection (Leung et al. 2002), methods with high sensitivity are needed to increase the likelihood of successful management outcomes. One such method is the isolation of DNA from environmental samples (environmental DNA or eDNA) such as water or sediment for the detection of organisms. Studies have demonstrated that the amplification of DNA barcode regions from eDNA (i.e. eDNA metabarcoding) can be used to detect marine NNS (Grey et al. 2018, Holman et al. 2019a, Rey et al. 2020, Duarte et al. 2021) and that it is a sensitive and accurate method for biomonitoring (Deiner et al. 2017, Fediajevaite et al. 2021). However, eDNA surveys are rarely used in conjunction with existing methods to detect NNS range shifts, and eDNA metabarcoding can validate, endorse, or highlight flaws, in current biodiversity management strategies.

During a range expansion, understanding if there was a single NNS introduction event or multiple simultaneous introductions is valuable for managers to target possible source regions, and to effectively manage introduction vectors (e.g. ballast waters). As NNS spread across the new region, understanding if expansions are due to local spread or introductions from distant regions is useful to target containment efforts. Finally, after eradication efforts have been conducted, understanding if the reappearance of NNS is due to incomplete eradication or a secondary reintroduction is of value for effective management into the future. The sequencing of DNA isolated from NNS has previously identified the source of NNS (Brown and Stepien 2009, Hudson et al. 2020), provided evidence of multiple introductions (Jeffery et al. 2017) and tested if post eradication invasions are a result of incomplete eradiation or reinvasion (Russell et al. 2010). Cumulatively, these studies have demonstrated the value of DNA evidence for the management of NNS. Furthermore, observations from both laboratory (Holman et al. 2019b, Tsuji et al. 2020) and field studies (Sigsgaard et al. 2017, Adams et al. 2019, Sigsgaard et al. 2020, Turon et al. 2020, Andres et al. 2021, Weitemier et al. 2021) have shown that eDNA can provide population genetics inference, but very little work has used this approach to study NNS (Uchii et al. 2016).

Here we combined eDNA metabarcoding, mitochondrial gene sequencing, and non-molecular biodiversity surveys to study four NNS that are directly relevant to marine natural resource managers. First, we evaluated if the NNS shifted their ranges over decadal time scales and compared each range shift to historical data. Secondly, we evaluated changes in genetic diversity and haplotype composition for each NNS between two sampling occasions across the sampled coastline. Finally, we examined how spatial genetic variation data can inform the management of range shifting species by comparing eDNA metabarcoding data to biodiversity survey and mitochondrial DNA sequence datasets.

4.4 Methods

4.4.1 Fieldwork and historical biodiversity data

The coastline of South Africa is an ideal system to study range shifting species and their management. South Africa has been subject to intense human impact and many species invasions have been documented across the three environmentally varied coastal ecoregions (Robinson et al. 2005, Griffiths et al. 2010, Rius et al. 2014a). Moreover, rapid assessment survey (a non-molecular biodiversity survey technique) (Arenas et al. 2006) data has been previously collected and mitochondrial sequence data have been generated for NNS along the entire coastline (Rius et al. 2014a). Furthermore, historical data are available for a range of relevant species (Michaelsen and Stephenson 1934, Millar 1962, Monniot et al. 2001) providing an insightful opportunity to conduct a spatial and temporal analysis of range expansions. Here, we selected twelve human impacted sites and conducted surveys (see details below) between October and November 2017. The sampled sites were the 11 sites previously sampled in 2007 and 2009 by Rius et al. (2014a), which included all major harbours and a number of marinas, and a new marina constructed post 2009 (Figure 1a with full details in Supplementary Table 1). Collectively, the sites encompass the main introduction points for marine NNS into the South African coastline.

At each sampling site a rapid assessment survey was conducted following Rius et al. (2014a), targeting non-indigenous ascidian species (Class: Ascidiacea). Ascidians are unique species for studying range expansions as they are successful invaders (Zhan et al. 2015) and have a relatively short pelagic larval phase, meaning that long-distance dispersal can only be achieved through anthropogenic transport of species (Svane and Young 1989). For each site, species abundance was ranked as absent (0%), scarce (< 10%), common (10-50%), or dominant (> 50%) based on observations of substrate coverage as in (Rius et al. 2014a).

Rapid assessment survey data from 2007 and 2009 was sourced from Rius et al. (2014a) for the species of interest. Additionally, historical incidence data was extracted from several taxonomic publications (Michaelsen and Stephenson 1934, Millar 1955, Millar 1962, Millar

1964, Monniot et al. 2001). These investigations are not an exhaustive survey of the coastline, but they provide valuable historical species incidence data over the last century and are therefore of value in gaining a broad understanding of range shifts over time.

4.4.2 Sample collection, DNA extraction and sequencing

Tissue samples were collected for species for which genetic data were available from the 2009 surveys (*Ciona intestinalis, Clavelina lepadiformis, Microcosmus squamiger* and *Styela plicata*). Samples were collected where sufficient numbers of individuals per species were present at a site to provide a reasonable estimate of genetic diversity (minimum 10 individuals), with 30+ individuals per species being the target at each sampling site. Organisms were sampled by hand, with no adjacent (within 0.3m) individuals collected, and dissected within six hours (see details of research permit in the Acknowledgements). For each sampled individual, approximately 10mm^2 of tissue from around the syphons was dissected using tools decontaminated with 10% bleach solution (3.5% chlorine), except in the case of *C. lepadiformis* for which a single zooid was removed from the tunic and stored. Tissue samples were preserved in 100% ethanol and stored at ambient temperature during transportation, and then stored at -80°C in the laboratory until later DNA extraction.

DNA from ascidian tissue samples was extracted using the Qiagen (Hilden, Germany) DNeasy Blood and Tissue Kit (96 Well Format) following manufacturer's recommended protocol with one blank control per extraction run. The final DNA elution was performed using 200μl of Qiagen Buffer ATL. A section of the cytochrome c oxidase subunit I gene (COI) was sequenced for all tissue samples aiming to cover the entire section previously analysed in Rius et al. (2014a). Each PCR contained 6μl of Applied Biosystems (Foster City, Califonia, USA) AmpliTaq GOLD 360 Mastermix, 1.8μl of oligonucleotide mix (5 μm concentration per primer), 1.2μl of undiluted template DNA and PCR quality water up to 12μl total reaction volume. The reaction conditions varied by primer set and are listed in Supplementary Table 2a. During preliminary trials a set of primers were designed and validated for *M. squamiger* (sequence details in Supplementary Table 2b), existing primer sets

(Tarjuelo et al. 2001, Nydam and Harrison 2007, Steinke et al. 2016) were optimised for the remaining three species. Successful amplification was confirmed using gel electrophoresis and PCR products were cleaned using Applied Biosystems ExoSAP-IT Express following the manufacturer's recommended protocol. Cleaned products were normalised to approximately 50ng/μl and 5μl of sample was added to each of 5μl of the forward or reverse primers (5μm) used in the initial PCR. These samples were sent for sequencing using the Macrogen Europe (Amsterdam, Netherlands) EZ-Seq service. Resultant chromatogram files were analysed using Geneious Prime (v2020.2.4) (Biomatters Ltd, Auckland, New Zealand). For each sequence the forward and reverse traces were aligned and sequences with ambiguities or failed reactions were re-sequenced from the initial PCR once and subsequently discarded if poor results persisted. The 764 COI sequences from Rius et al. (2014a) were added to the analysis and trimmed, truncated and aligned with the experimental data as follows. For each species, sequences were trimmed to remove primer binding and poor-quality regions and aligned using the Geneious Alignment Tool. Subsequently each alignment was manually checked to confirm complete alignment, and short sequences that did not overlap at all polymorphic regions or had ambiguous base calls were discarded.

4.4.3 Environmental DNA metabarcoding

Before each rapid assessment survey, surface seawater was sampled from the top 10cm for eDNA metabarcoding following Holman et al. (2021b). Briefly, three replicate 400ml water samples were filtered on site with a 0.22 μm polyethersulfone enclosed filter. Filters were preserved with Longmire's solution until DNA extraction following Spens et al. (2017). Data generated from these samples is presented in Holman et al. (2021b) with the aim of conservatively characterising whole community diversity. COI and ribosomal RNA (18S) data targeting metazoans (Leray et al. 2013, Zhan et al. 2013) was reanalysed as follows for accurate ascidian species detection. Primer regions were removed from forward and reverse reads using the default settings of Cutadapt (v2.3) (Martin 2011). Sequences were denoised and an ASV (amplicon sequence variant) by sample table generated using DADA2 (v1.12) (Callahan et al. 2016) in R (v3.6.1) (R_Core_Team 2021) with parameters as in Holman et al.

(2021b). Recent work has highlighted that different bioinformatic methods have an effect on the resolution of intra-specific variation of eDNA metabarcoding data (Tsuji et al. 2020, Turon et al. 2020, Antich et al. 2021). Therefore in addition to the sequenced tissue samples and DADA2 methods outlined above, we reanalysed the COI data using the *unoise3* algorithm (Edgar 2013) as follows. Raw COI paired-end fastq data from Holman et al. (2021b) was merged using usearch (v11.0.667) (Edgar 2013) with the following parameters -*fastq_maxdiffs 15 -fastq_pctid 80*. Primer sequences were then stripped from each merged read using Cutadapt (v3.1) (Martin 2011) under the default parameters and reads longer than 323 and shorter than 303 base pairs (±10 from the expected size of 313) were discarded. Reads from all samples were pooled, and singletons and reads with an expected error greater than 1 were discarded using vsearch (v2.15.1) (Rognes et al. 2016). The *unoise3* algorithm from usearch was then used to generate ASVs with -*unoise_alpha* set at 5 as recommended for resolving metazoan intraspecific variation with a COI fragment of 313 base pairs in length (Turon et al. 2020). Sequences were then mapped back to the ASVs using the -u*search_global* function of vsearch with an -*id* parameter of 0.995 to produce an ASV by sample table.
To provide an initial taxonomic assignment all ASVs were compared using a BLAST (v2.6.0+) search with no limits on sequence similarity or match length to the NCBI *nt* database (downloaded 16th May 2019). Taxonomic assignments were then parsed using a custom R function (*ParseTaxonomy*, DOI:10.5281/zenodo.4671710) with the default settings. The taxonomic assignments were subset to include only those with a hit to species in the class Ascidiacea. The following quality control steps were then applied to each dataset. The data was filtered to only retain ASVs that appeared in more than one replicate sample. For any ASVs detected in both the negative and experimental control samples, the maximum number of reads in the negative controls were subtracted from the experimental control samples. Reads were then divided by the total number of reads per sample and relative proportions were used in all subsequent analyses, technical replicates per site were averaged. The remaining ASVs were then taxonomically checked manually using the online National Centre for Biotechnology nucleotide BLAST search function against the *nt* databases (last accessed on 1st October 2020) under default megablast parameters. For each ASV in the COI dataset, taxonomy was only assigned at species level if multiple, independent sequences had a

match greater than 97% identity (with 100% coverage) with no other species within 97% of the target ASV. For the 18S dataset a 100% match (with 100% coverage) between the subject ASV and database sequences was required for taxonomic confirmation. Additionally, as some taxa within the same genera have near 100% similarity at the 18S region, taxonomy was only assigned to species if organisms from the same genera were in the database with at least 1 base pair between the query and species from the same genera. Following taxonomic annotation, ASVs assigned to the same species were merged for the distribution datasets. ASVs were kept separate for the haplotype reconstruction of the COI data.

4.4.4 Data manipulation and statistical analysis

Distances between sites along the coast were estimated by drawing a transect 1 km parallel to the coastline in Google Earth Pro (v7.3.2.5776) and calculating the distance between each pair of sites. The study area was plotted using the function *map* from the package *maps* (v3.3.0). Sequenced COI regions from 2009 and 2017 were aligned separately for each species using the Geneious aligner in Geneious Prime; alignments were truncated to include only overlapping regions. Sequences were manipulated using the *SeqinR* package in R (v4.2-5) (Charif and Lobry 2007). Nucleotide and haplotypic diversity were calculated using the *nuc.div* and *hap.div* functions from the *pegas* package (v0.14) (Paradis 2010). For each species an alignment was created between the tissue sampled COI sequences and the eDNA metabarcoding derived haplotypes. The region of overlap was extracted and used in subsequent analyses. Haplotype frequencies were calculated per site for the tissue derived sequences and the different bioinformatic analyses of eDNA metabarcoding data. Minimum spanning network haplotype maps (Bandelt et al. 1999) were created using the default settings of PopArt (v1.7) (Leigh and Bryant 2015). Analyses of molecular variance (AMOVA) were performed using the function *poppr.amova* from the *poppr* package (v2.8.6) (Kamvar et al. 2014). AMOVA models were structured to analyse the effect of sampling year and sites for each species. All data analyses were conducted in R (v4.0.3) unless otherwise stated.

4.5 Results

4.5.1 Range shifts

Rapid assessment surveys found that non-native ascidians known to be broadly restricted to warmer waters (*M. squamiger* and *S. plicata*) (Rius et al. 2014a) showed distributions principally limited to the southern and eastern coastlines (Fig.1b). In contrast *C. robusta* and *C. lepadiformis* were found along most of the coastline. We found no change across years in range extent for *M. squamiger* and *S. plicata*, a decrease in easternly range for *C. robusta* and an expansion of range both westerly and easterly for *C. lepadiformis* (Fig.1c). Historical total range extent data (Fig.1d) showed more recent increases in range for *C. lepadiformis* and *S. plicata* compared to *C. robusta* and *M. squamiger*. The COI and 18S eDNA metabarcoding data showed mixed results. There was good agreement between detections from eDNA and rapid assessment surveys in *M. squamiger* and *S. plicata* (see Fig.1b). However, 18S entirely failed to detect *C. robusta* or *C. lepadiformis*, and COI demonstrated a number of false-negative metabarcoding detections in these species (Fig.1b). For sites sharing detections from eDNA metabarcoding and rapid assessment surveys, eDNA metabarcoding data and field density estimates showed a non-significant relationship (18S $p = 0.052$, COI $p = 0.297$) (see Supplementary Note 1 for details).

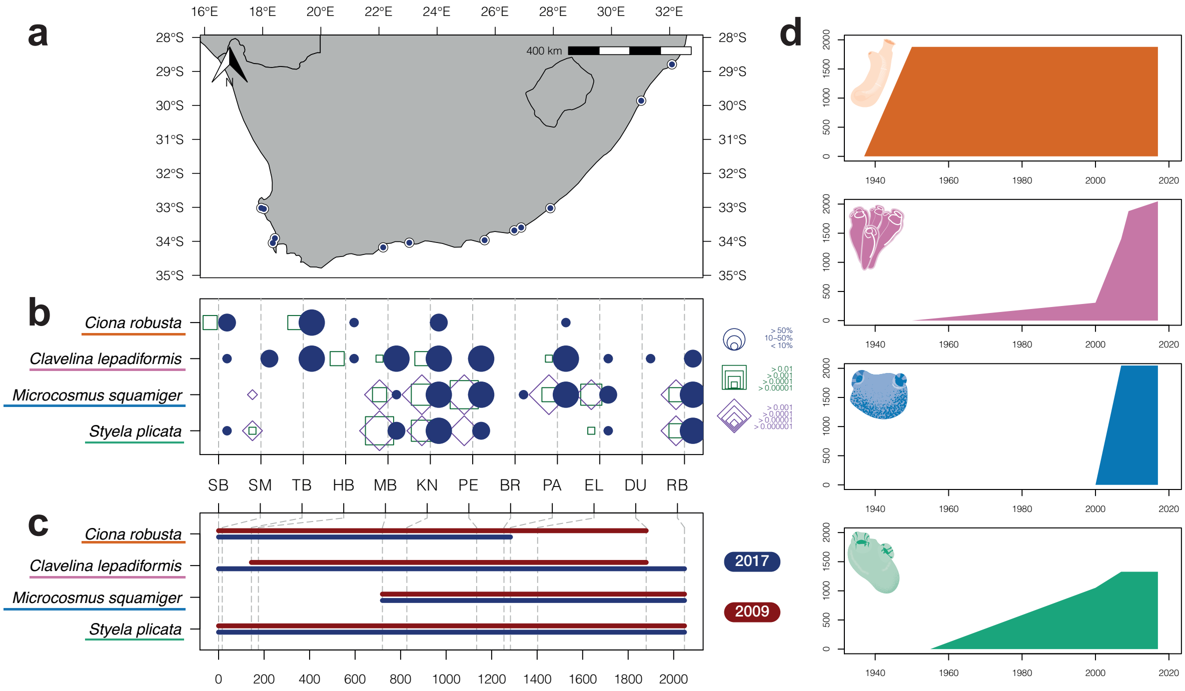


Figure 1 a Map depicting the coastline of South Africa, sampling sites are shown as blue points, full details in Supplementary Table 1. **b** Bubble plot showing incidence of four non-native ascidians across the sampling sites shown in the map from west to east. Blue bubbles show results of rapid assessment surveys and square outlines show the results of eDNA metabarcoding surveys conducted concurrently. Results from COI are shown with green squares and 18S shown with purple squares, the size of each point or square shows the comparative density. Site codes correspond with sites as detailed in Supplementary Table 1. **c** Line plot showing range extent over the surveyed coast for 2009 (dark red) rapid assessment surveys from Rius et al. (2014a) and surveys conducted in 2017 presented here (blue). The location of each site across the coastline is shown with grey dashed lines. **d** Historical maximum range extent for each of the featured species across the coastline of South Africa, y axis is kilometres of extent, x axis is year, colour indicates each of the species indicated according to labels in b and c.

4.5.2 Changes in genetic composition

A total of 1,320 sequencing reactions generated 660 bi-directionally sequenced COI sequences. After alignment and quality control, 541 samples remained with complete alignment and no missing site information, 88 for *C. robusta,* 261 for *C. lepadiformis,* 90 for *M. squamiger* and 102 for *S. plicata*. After combining the COI sequences with previously sequenced samples from 2009 (Rius et al. 2014a), alignments were 626, 440, 635 and 599 base pairs in length for *C. robusta*, *C. lepadiformis*, *M. squamiger* and *S. plicata* respectively.

Observed haplotype richness across both sampling years and all sites was highest in *M. squamiger* followed by *C. robusta, S. plicata and C. lepadiformis* (Fig. 2). There was no statistically significant difference between nucleotide or haplotype diversity between sampling years across all species ($p > 0.05$ in all cases, see Supplementary Note 2 for full model output and details). Additionally, AMOVA models found no significant differences between sampling years across all species ($p > 0.05$ in all cases, see Supplementary Table 3 for full model outputs), but significant differences between sampling sites within years ($p < 0.05$ in all species, see Supplementary Table 3 for full model outputs). In all species, the greatest proportion of the genetic variance was found between samples, then within sampling sites, followed by variance between sampling sites (Supplementary Table 3). As shown in Figure 2, haplotype frequencies agreed with the AMOVA analyses, showing stable patterns of genetic variation occurring between years and variation in haplotype frequencies across the study system (Fig. 2).

After aligning the shorter sequences derived from eDNA metabarcoding data to the sequenced COI region, alignments were 191, 258, 289 and 286 base pairs in length for *C. robusta*, *C. lepadiformis*, *M. squamiger* and *S. plicata* respectively. Regardless of bioinformatic method and across species, the eDNA metabarcoding data did not recover all the haplotype sequences derived from tissue (Figure 3).

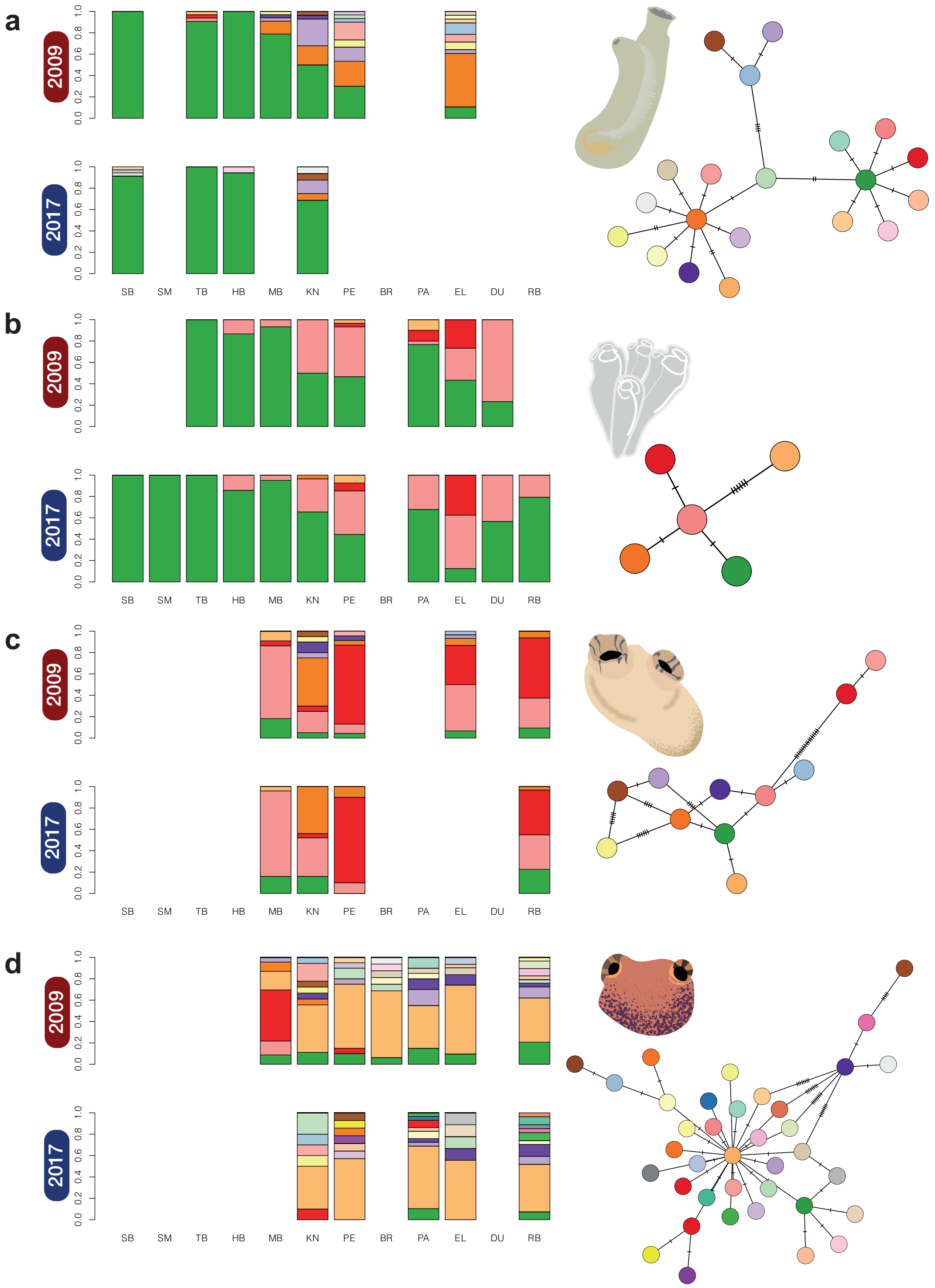


Figure 2 Mitochondrial DNA COI haplotype proportions for **a** *Ciona robusta* **b** *Clavelina lepadiformis* **c** *Styela plicata* **d** *Microcosmus squamiger* across the South African coastline. Results are shown for surveys conducted in 2009 and 2017 for each species, site abbreviations follow supplementary Table 1. Haplotype networks based on minimum spanning distance are shown for each species with colours matching the bar plot within species, the number of cross-hatches indicates the mutation steps between haplotypes.

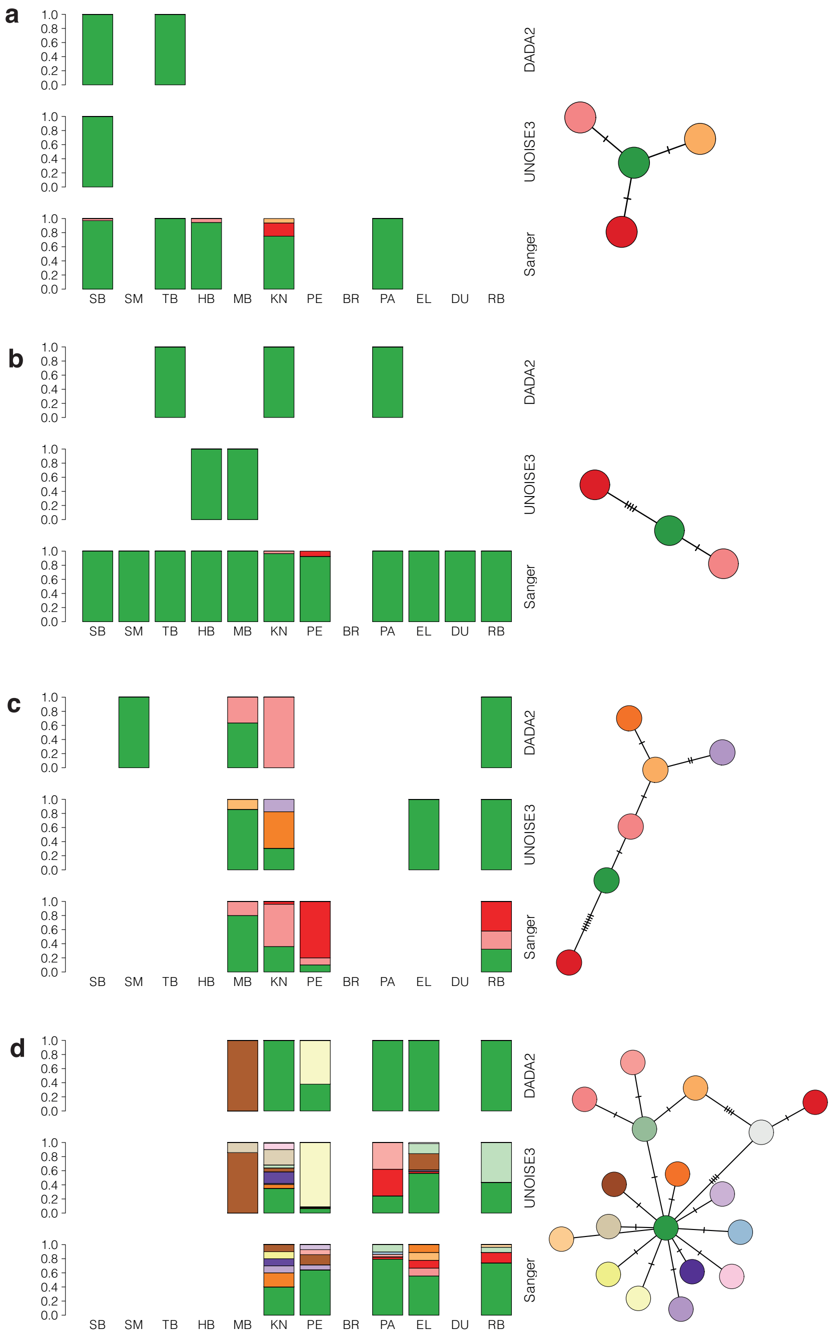


Figure 3 Haplotype proportions recovered using eDNA metabarcoding for **a** *Ciona robusta* **b** *Clavelina lepadiformis* **c** *Styela plicata* **d** *Microcosmus squamiger* across the South African coastline. Results are shown for analysis of COI eDNA metabarcoding data using the denoising software DADA2 and UNOISE3 for each species, site abbreviations follow supplementary Table 1. Haplotype networks based on minimum spanning distance are shown for each species with colours matching the bar plot within species, the number of cross-hatches indicates single nucleotide mutation steps between haplotypes.

4.6 Discussion

Here we found both losses and gains in range size across sampling years for four non-native ascidian species, with no consistent pattern emerging when introduction dates were compared. For all species we found haplotype variability across the study region but no significant change in genetic variation for almost a decade. Finally, eDNA metabarcoding data recovered broad NSS incidence trends and for some species was as accurate as non-molecular surveys. Most dominant haplotypes from tissue samples were detected with eDNA metabarcoding but fine scale genetic patterns could not be resolved using the eDNA metabarcoding data. Cumulatively, the evidence demonstrates that a suite of tools, including DNA and non-DNA biodiversity surveys can be used in combination to evaluate the role of genetic variation on range shifts and to inform natural resource managers.

Non-DNA biodiversity surveys found that *C. lepadiformis* expanded its range by 168.4 km since surveys in 2009, for an assumed rate of 21.1 km per year. This is in line with previous studies that found an average marine non-native spread rate of 44.3 km per year (Sorte et al. 2010), with values of 16 km per year for tunicates, 30.0 km per year for barnacles and 20 km per year for a bryozoan species (Grosholz 1996). In contrast, we observed a range contraction for *C. robusta*, which was unexpected. There are limited studies showing range contraction in the introduced range for marine species. However, previous work has identified biotic resistance for invasions of several species in the genus *Ciona* (Dumont et al. 2011, Rius et al. 2014b), and so it might be feasible for local species to have begun predating on *Ciona robusta* during the 80+ years it has been documented in South Africa (Fig.1d). A lack of any western increase in range for *M. squamiger* might be explained by the species inability to mature to reproductive age in the colder sea temperature on the western coast (Rius et al. 2014a). Further range expansions or contractions (eastwards for *M. squamiger* and east or westward for *S. plicata*) cannot be ruled out as observations of these species extended to the margins of the sampled area. It is important to note that the harbours and marinas in this study act as islands of suitable habitat, and the frequency of introductions outside these areas is relatively uncertain. Further surveys of surrounding hard benthic environments are required to understand the role of artificial environments across the coastal ecosystem. Overall, these patterns demonstrate that the spread of marine non-native species is not characterised by a continuous expansion of range, but rather by a complex picture of expansions and contractions in response to dynamic abiotic and biotic conditions.

A consistent pattern of genetic differentiation emerged across the studied species; significant differences across sampling sites and persistence of similar haplotypes across time (Fig. 2). Previous studies of temporal changes in genetic diversity of non-native ascidians have found some evidence for genetic differences over time (Pérez-Portela et al. 2012, Pineda et al. 2016b). In contrast other work has found relatively stable genetic diversity over several years (Pineda et al. 2016a, Haye et al. 2021). In our study, the time between sampling occasions (i.e. 2009 and 2017) represents between four and 24 generations depending on the species (De Caralt et al. 2002, Rius et al. 2009, Pineda et al. 2013). Therefore, dramatic changes in haplotype frequencies could only be as a result of anthropogenic transfer of haplotypes between sites or changes in site frequencies in response to high mortality events (for example extreme weather events). These types of changes have been documented in ascidian species elsewhere (Pérez-Portela et al. 2012, Chang et al. 2018, Caputi et al. 2019), and a large number of NNS introductions have been documented in South African marinas and harbours supporting regional transfer of these organisms (Robinson et al. 2005, Rius et al. 2014a). It is therefore somewhat surprising that across four different species, all of which are known to be transported anthropogenically, there was little evidence of shifts in haplotype composition. Consequently, our results demonstrated that the studied NSS are well-established and are not subject to high levels of mortality or genetic bottleneck that may affect population viability. It may be that these well-established haplotypes impede newcomers to succeed and ultimately change the haplotype composition of the site.

We found that eDNA metabarcoding captured similar incidence data as rapid assessment surveys for some species, and performed poorly for others. Previous work has identified that NNS can be detected using eDNA metabarcoding (Holman et al. 2019a, Rey et al. 2020), but these surveys aimed at detecting any NNS rather than a specific set of target taxa. Several studies have identified that general target metabarcoding primers show lower reliability and sensitivity compared to species-specific quantitative PCR assays (Harper et al. 2018, Blackman et al. 2020). Additionally, previous work has identified that in some cases different bioinformatic methods carry variable sensitivity (Scott et al. 2018), although this effect is fairly minimal in this dataset (see Supplementary Note 3). Managers should therefore be aware that general metabarcoding primers will perform well in the detection of some important NNS but others may be missed due to poor sensitivity. In cases when a list of priority species can be assembled, mixed DNA positive control samples or trials with aquaria of known composition (for example Holman et al. (2021a)) would provide information on which NNS might be overlooked by eDNA metabarcoding. Inevitably, there will be a cost-benefit trade-off between using imperfect broad-target metabarcoding assays for monitoring unknown invaders, and expending resources on the development and application of eDNA tools targeting specific known NNS.

In some cases, resource managers might be interested in tracking invasions using haplotype data (Darling 2015). Here, we showed that eDNA metabarcoding with broad-target primers resolves broad scale patterns of haplotype diversity (Fig. 3). However, fine-scale genetic variation was not recovered in our study, indicating that targeted eDNA amplicon sequencing (Sigsgaard et al. 2017) might be more appropriate when this level of genetic data is required. As with biodiversity incidence data, the management objectives for a given NNS determine how haplotype sequencing should be implemented. If large numbers of tissue samples can be easily collected and there are sufficient resources, then sequencing the tissue directly might be more appropriate. In contrast if a broad scale analysis across a larger or difficult-to-sample area is preferred resolving haplotype data from eDNA metabarcoding might be preferable. Overall, eDNA based techniques show great potential for NNS detection but for our target taxa, we demonstrated that current biodiversity surveys and direct tissue sequencing are more reliable for the detection of NNS and genetic composition. It is important to note that there are several key advantages of eDNA-based methods compared to the other tools used in this work. Firstly, eDNA samples can be collected with minimal training and the sequenced DNA provides an unambiguous identification, provided reference data is available (Grey et al. 2018, Rey et al. 2020). Secondly, eDNA-based methods can be automated and can scale to a much greater survey effort at reduced cost compared to other methods (Gold et al. 2021). Third, the limitations described above concerning the sensitivity of eDNA-based incidence data and lack of resolution of eDNA-based haplotype data can be attributed to the use of metabarcoding with broad-target primers. Reanalysing the samples with metabarcoding primers for more specific groups or using species specific qPCR assays (Harper et al. 2018) would provide increased sensitivity and accuracy.

Overall we demonstrated how our combined methodological approach can effectively detect spatial and temporal trends of range shifts and genetic differentiation, but also to monitor biodiversity changes of both threatened and NSS. The strengths of eDNA or DNA-based biomonitoring demonstrated here for the detection of range shifting species make them a pragmatic choice for natural resources managers. These tools provide managers with greater sensitivity and accuracy when monitoring biodiversity in human impacted environments.

4.7 References

Adams, C. I., M. Knapp, N. J. Gemmell, G.-J. Jeunen, M. Bunce, M. D. Lamare, and H. R. Taylor. 2019. Beyond biodiversity: can environmental DNA (eDNA) cut it as a population genetics tool? Genes **10**:192.

Airoldi, L., X. Turon, S. Perkol-Finkel, and M. Rius. 2015. Corridors for aliens but not for natives: effects of marine urban sprawl at a regional scale. Diversity and Distributions **21**:755-768.

Andres, K. J., S. A. Sethi, D. M. Lodge, and J. Andrés. 2021. Nuclear eDNA estimates population allele frequencies and abundance in experimental mesocosms and field samples. Molecular Ecology **30**:685-697.

Antich, A., C. Palacin, O. S. Wangensteen, and X. Turon. 2021. To denoise or to cluster, that is not the question: optimizing pipelines for COI metabarcoding and metaphylogeography. BMC Bioinformatics **22**:177.

Arenas, F., J. D. D. Bishop, J. T. Carlton, P. J. Dyrynda, W. F. Farnham, D. J. Gonzalez, M. W. Jacobs, C. Lambert, G. Lambert, S. E. Nielsen, J. A. Pederson, J. S. Porter, S. Ward, and C. A. Wood. 2006. Alien species and other notable records from a rapid assessment survey of marinas on the south coast of England. Journal of the Marine Biological Association of the United Kingdom **86**:1329-1337.

Bandelt, H.-J., P. Forster, and A. Röhl. 1999. Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution **16**:37-48.

Bax, N., A. Williamson, M. Aguero, E. Gonzalez, and W. Geeves. 2003. Marine invasive alien species: a threat to global biodiversity. Marine Policy **27**:313-323.

Blackman, R. C., K. K. S. Ling, L. R. Harper, P. Shum, B. Hanfling, and L. Lawson-Handley. 2020. Targeted and passive environmental DNA approaches outperform established methods for detection of quagga mussels, *Dreissena rostriformis bugensis* in flowing water. Ecology and Evolution **10**:13248-13259.

Booy, O., P. A. Robertson, N. Moore, J. Ward, H. E. Roy, T. Adriaens, R. Shaw, J. Van Valkenburg, G. Wyn, S. Bertolino, O. Blight, E. Branquart, G. Brundu, J. Caffrey, D. Capizzi, J. Casaer, O. De Clerck, N. E. Coughlan, E. Davis, J. T. A. Dick, F. Essl, G. Fried, P. Genovesi, P. González-Moreno, F. Huysentruyt, S. R. Jenkins, F. Kerckhof, F. E. Lucy, W. Nentwig, J. Newman, W. Rabitsch, S. Roy, U. Starfinger, P. D. Stebbing, J. Stuyck, M. Sutton-Croft, E. Tricarico, S. Vanderhoeven, H. Verreycken, and A. C. Mill. 2020. Using structured eradication feasibility assessment to prioritize the management of new and emerging invasive alien species in Europe. Global Change Biology **26**:6235-6250.

Brown, J. E., and C. A. Stepien. 2009. Invasion genetics of the Eurasian round goby in North America: tracing sources and spread patterns. Molecular Ecology **18**:64-79.

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. Johnson, and S. P. Holmes. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods **13**:581-583.

Caputi, L., F. Toscano, M. Arienzo, L. Ferrara, G. Procaccini, and P. Sordino. 2019. Temporal correlation of population composition and environmental variables in the marine invader Ciona robusta. Marine Ecology **40**:e12543.

Cardinale, B. J., J. E. Duffy, A. Gonzalez, D. U. Hooper, C. Perrings, P. Venail, A. Narwani, G. M. Mace, D. Tilman, D. A. Wardle, A. P. Kinzig, G. C. Daily, M. Loreau, J. B. Grace, A. Larigauderie, D. S. Srivastava, and S. Naeem. 2012. Biodiversity loss and its impact on humanity. Nature **486**:59-67.

Chang, A. L., C. W. Brown, J. A. Crooks, and G. M. Ruiz. 2018. Dry and wet periods drive rapid shifts in community assembly in an estuarine ecosystem. Global Change Biology **24**:e627-e642.

Charif, D., and J. R. Lobry. 2007. SeqinR 1.0-2: a contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. Pages 207-232 Structural approaches to sequence evolution. Springer.

Clout, M. N., and P. A. Williams. 2009. Invasive species management: a handbook of principles and techniques. Oxford University Press.

Darling, J. A. 2015. Genetic studies of aquatic biological invasions: closing the gap between research and management. Biological Invasions **17**:951-971.

Darling, J. A., B. S. Galil, G. R. Carvalho, M. Rius, F. Viard, and S. Piraino. 2017. Recommendations for developing and applying genetic tools to assess and manage biological invasions in marine ecosystems. Marine Policy **85**:54-64.

De Caralt, S., S. López-Legentil, I. Tarjuelo, M. J. Uriz, and X. Turon. 2002. Contrasting biological traits of Clavelina lepadiformis (Ascidiacea) populations from inside and outside harbours in the western Mediterranean. Marine Ecology Progress Series **244**:125-137.

Deiner, K., H. M. Bik, E. Machler, M. Seymour, A. Lacoursiere-Roussel, F. Altermatt, S. Creer, I. Bista, D. M. Lodge, N. de Vere, M. E. Pfrender, and L. Bernatchez. 2017. Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. Molecular Ecology **26**:5872-5895.

Diagne, C., B. Leroy, A.-C. Vaissière, R. E. Gozlan, D. Roiz, I. Jarić, J.-M. Salles, C. J. A. Bradshaw, and F. Courchamp. 2021. High and rising economic costs of biological invasions worldwide. Nature **592**:571-576.

Dornelas, M., N. J. Gotelli, H. Shimadzu, F. Moyes, A. E. Magurran, and B. J. McGill. 2019. A balance of winners and losers in the Anthropocene. Ecology Letters **22**:847-854.

Duarte, S., P. E. Vieira, A. S. Lavrador, and F. O. Costa. 2021. Status and prospects of marine NIS detection and monitoring through (e)DNA metabarcoding. Science of the Total Environment **751**:141729.

Dumont, C. P., C. F. Gaymer, and M. Thiel. 2011. Predation contributes to invasion resistance of benthic communities against the non-indigenous tunicate *Ciona intestinalis*. Biological Invasions **13**:2023-2034.

Edgar, R. C. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nature Methods **10**:996-998.

Fediajevaite, J., V. Priestley, R. Arnold, and V. Savolainen. 2021. Meta-analysis shows that environmental DNA outperforms traditional surveys, but warrants better reporting standards. Ecology and Evolution **11**:4803-4815.

Gamfeldt, L., J. S. Lefcheck, J. E. K. Byrnes, B. J. Cardinale, J. E. Duffy, and J. N. Griffin. 2015. Marine biodiversity and ecosystem functioning: what's known and what's next? Oikos **124**:252-265.

Giakoumi, S., S. Katsanevakis, P. G. Albano, E. Azzurro, A. C. Cardoso, E. Cebrian, A. Deidun, D. Edelist, P. Francour, C. Jimenez, V. Mačić, A. Occhipinti-Ambrogi, G. Rilov, and Y. R. Sghaier. 2019. Management priorities for marine invasive species. Science of the Total Environment **688**:976-982.

Glasby, T. M., S. D. Connell, M. G. Holloway, and C. L. Hewitt. 2007. Nonindigenous biota on artificial structures: could habitat creation facilitate biological invasions? Marine Biology **151**:887-895.

Gold, Z., J. Sprague, D. J. Kushner, E. Zerecero, and P. H. Barber. 2021. eDNA metabarcoding as a biomonitoring tool for marine protected areas. PLoS One **16(2)**:e0238557.

Grey, E. K., L. Bernatchez, P. Cassey, K. Deiner, M. Deveney, K. L. Howland, A. Lacoursiere-Roussel, S. C. Y. Leong, Y. Li, B. Olds, M. E. Pfrender, T. A. A. Prowse, M. A. Renshaw, and D. M. Lodge. 2018. Effects of sampling effort on biodiversity patterns estimated from environmental DNA metabarcoding surveys. Scientific Reports **8**:8843.

Griffiths, C. L., T. B. Robinson, L. Lange, and A. Mead. 2010. Marine biodiversity in South Africa: an evaluation of current states of knowledge. PLoS One **5**:e12008.

Grosholz, E. D. 1996. Contrasting Rates of Spread for Introduced Species in Terrestrial and Marine Systems. Ecology **77**:1680-1686.

Harper, L. R., L. Lawson Handley, C. Hahn, N. Boonham, H. C. Rees, K. C. Gough, E. Lewis, I. P. Adams, P. Brotherton, S. Phillips, and B. Hänfling. 2018. Needle in a haystack? A comparison of eDNA metabarcoding and targeted qPCR for detection of the great crested newt (*Triturus cristatus*). Ecology and Evolution **8**:6330-6341.

Haye, P. A., X. Turon, and N. I. Segovia. 2021. Time or Space? Relative Importance of Geographic Distribution and Interannual Variation in Three Lineages of the Ascidian Pyura chilensis in the Southeast Pacific Coast. Frontiers in Marine Science **8**:413.

Holman, L. E., Y. Chng, and M. Rius. 2021a. How does eDNA decay affect metabarcoding experiments? Environmental DNA **In Press**.

Holman, L. E., M. de Bruyn, S. Creer, G. Carvalho, J. Robidart, and M. Rius. 2019a. Detection of introduced and resident marine species using environmental DNA metabarcoding of sediment and water. Scientific Reports **9**:11559.

Holman, L. E., M. de Bruyn, S. Creer, G. Carvalho, J. Robidart, and M. Rius. 2021b. Animals, protists and bacteria share marine biogeographic patterns. Nature Ecology & Evolution **5**:738-746.

Holman, L. E., C. M. Hollenbeck, T. J. Ashton, and I. A. Johnston. 2019b. Demonstration of the Use of Environmental DNA for the Non-Invasive Genotyping of a Bivalve Mollusk, the European Flat Oyster (*Ostrea edulis*). Frontiers in Genetics **10**:1159.

Hudson, J., J. C. Castilla, P. R. Teske, L. B. Beheregaray, I. D. Haigh, C. D. McQuaid, and M. Rius. 2021. Genomics-informed models reveal extensive stretches of coastline under threat by an ecologically dominant invasive species. Proceedings of the National Academy of Sciences **118**:e2022169118.

Hudson, J., K. Johannesson, C. D. McQuaid, and M. Rius. 2020. Secondary contacts and genetic admixture shape colonization by an amphiatlantic epibenthic invertebrate. Evolutionary Applications **13**:600-612.

Jeffery, N. W., C. DiBacco, M. Van Wyngaarden, L. C. Hamilton, R. R. E. Stanley, R. Bernier, J. FitzGerald, K. Matheson, C. H. McKenzie, P. Nadukkalam Ravindran, R. Beiko, and I. R. Bradbury. 2017. RAD sequencing reveals genomewide divergence between independent invasions of the European green crab (*Carcinus maenas*) in the Northwest Atlantic. Ecology and Evolution **7**:2513-2524.

Kamvar, Z. N., J. F. Tabima, and N. J. Grünwald. 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. PeerJ **2**:e281.

Katsanevakis, S., A. Zenetos, C. Belchior, and A. C. Cardoso. 2013. Invading European Seas: Assessing pathways of introduction of marine aliens. Ocean & Coastal Management **76**:64-74.

Leigh, J. W., and D. Bryant. 2015. popart: full-feature software for haplotype network construction. Methods in Ecology and Evolution **6**:1110-1116.

Leray, M., J. Y. Yang, C. P. Meyer, S. C. Mills, N. Agudelo, V. Ranwez, J. T. Boehm, and R. J. Machida. 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.

Leung, B., D. M. Lodge, D. Finnoff, J. F. Shogren, M. A. Lewis, and G. Lamberti. 2002. An ounce of prevention or a pound of cure: bioeconomic risk analysis of invasive species. Proceedings of the Royal Society of London. Series B: Biological Sciences **269**:2407-2413.

Liordos, V., V. J. Kontsiotis, M. Georgari, K. Baltzi, and I. Baltzi. 2017. Public acceptance of management methods under different human–wildlife conflict scenarios. Science of the Total Environment **579**:685–693.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet Journal **17**:10–12.

Mazor, T., C. Doropoulos, F. Schwarzmueller, D. W. Gladish, N. Kumaran, K. Merkel, M. Di Marco, and V. Gagic. 2018. Global mismatch of policy and research on drivers of biodiversity loss. Nature Ecology & Evolution **2**:1071–1074.

Mazza, G., E. Tricarico, P. Genovesi, and F. Gherardi. 2014. Biological invaders are threats to human health: an overview. Ethology Ecology & Evolution **26**:112–129.

Michaelsen, W., and T. Stephenson. 1934. The ascidians of the Cape Province of South Africa. Transactions of the Royal Society of South Africa **22**:129–163.

Millar, R. 1955. A collection of ascidians from South Africa. Proceedings of the Zoological Society of London **125**:169–221.

Millar, R. 1964. South African ascidians collected by Th. Mortensen, with some additional material. Videnskabelige Meddelelser Fra Dansk Naturhistorisk Forening **127**:159–180.

Millar, R. H. 1962. Further descriptions of South African ascidians. South African Museum.

Molnar, J. L., R. L. Gamboa, C. Revenga, and M. D. Spalding. 2008. Assessing the global threat of invasive species to marine biodiversity. Frontiers in Ecology and the Environment **6**:485–492.

Monniot, C., F. Monniot, C. L. Griffiths, and M. Schleyer. 2001. South African ascidians. Annals of the South African Museum **108**:1–141.

Nydam, M. L., and R. G. Harrison. 2007. Genealogical relationships within and among shallow-water Ciona species (*Ascidiacea*). Marine Biology **151**:1839–1847.

Paradis, E. 2010. pegas: an R package for population genetics with an integrated–modular approach. Bioinformatics **26**:419-420.

Pecl, G. T., M. B. Araujo, J. D. Bell, J. Blanchard, T. C. Bonebrake, I. C. Chen, T. D. Clark, R. K. Colwell, F. Danielsen, B. Evengard, L. Falconi, S. Ferrier, S. Frusher, R. A. Garcia, R. B. Griffis, A. J. Hobday, C. Janion-Scheepers, M. A. Jarzyna, S. Jennings, J. Lenoir, H. I. Linnetved, V. Y. Martin, P. C. McCormack, J. McDonald, N. J. Mitchell, T. Mustonen, J. M. Pandolfi, N. Pettorelli, E. Popova, S. A. Robinson, B. R. Scheffers, J. D. Shaw, C. J. Sorte, J. M. Strugnell, J. M. Sunday, M. N. Tuanmu, A. Verges, C. Villanueva, T. Wernberg, E. Wapstra, and S. E. Williams. 2017. Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. Science **355**:6332.

Pérez-Portela, R., X. Turon, and J. Bishop. 2012. Bottlenecks and loss of genetic diversity: spatio-temporal patterns of genetic structure in an ascidian recently introduced in Europe. Marine Ecology Progress Series **451**:93–105.

Pineda, M.-C., B. Lorente, S. López-Legentil, C. Palacin, and X. Turon. 2016a. Stochasticity in space, persistence in time: genetic heterogeneity in harbour populations of the introduced ascidian *Styela plicata*. PeerJ **4**:e2158.

Pineda, M. C., S. López-Legentil, and X. Turon. 2013. Year-round reproduction in a seasonal sea: biological cycle of the introduced ascidian Styela plicata in the Western Mediterranean. Marine Biology **160**:221–230.

Pineda, M. C., X. Turon, R. Pérez-Portela, and S. López-Legentil. 2016b. Stable populations in unstable habitats: temporal genetic structure of the introduced ascidian *Styela plicata* in North Carolina. Marine Biology **163**:59.

Pluess, T., R. Cannon, V. Jarošík, J. Pergl, P. Pyšek, and S. Bacher. 2012. When are eradication campaigns successful? A test of common assumptions. Biological Invasions **14**:1365–1378.

R_Core_Team. 2021. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rey, A., O. C. Basurko, and N. Rodriguez-Ezpeleta. 2020. Considerations for metabarcoding-based port biological baseline surveys aimed at marine nonindigenous species monitoring and risk assessments. Ecology and Evolution **10**:2452-2465.

Rius, M., S. Clusella-Trullas, C. D. McQuaid, R. A. Navarro, C. L. Griffiths, C. A. Matthee, S. von der Heyden, and X. Turon. 2014a. Range expansions across ecoregions: interactions of climate change, physiology and genetic diversity. Global Ecology and Biogeography **23**:76–88.

Rius, M., M. C. Pineda, and X. Turon. 2009. Population dynamics and life cycle of the introduced ascidian Microcosmus squamiger in the Mediterranean Sea. Biological Invasions **11**:2181–2194.

Rius, M., E. E. Potter, J. D. Aguirre, and J. J. Stachowicz. 2014b. Mechanisms of biotic resistance across complex life cycles. Journal of Animal Ecology **83**:296–305.

Robinson, T. B., C. L. Griffiths, C. D. McQuaid, and M. Rius. 2005. Marine alien species of South Africa — status and impacts. African Journal of Marine Science **27**:297–306.

Rognes, T., T. Flouri, B. Nichols, C. Quince, and F. Mahe. 2016. VSEARCH: a versatile open source tool for metagenomics. PeerJ **4**:e2584.

Roy, H. E., S. Bacher, F. Essl, T. Adriaens, D. C. Aldridge, J. D. D. Bishop, T. M. Blackburn, E. Branquart, J. Brodie, C. Carboneras, E. J. Cottier-Cook, G. H. Copp, H. J. Dean, J. Eilenberg, B. Gallardo, M. Garcia, E. García-Berthou, P. Genovesi, P. E. Hulme, M. Kenis, F. Kerckhof, M. Kettunen, D. Minchin, W. Nentwig, A. Nieto, J. Pergl, O. L. Pescott, J. M. Peyton, C. Preda, A. Roques, S. L. Rorke, R. Scalera, S. Schindler, K. Schönrogge, J. Sewell, W. Solarz, A. J. A. Stewart, E. Tricarico, S. Vanderhoeven, G. van der Velde, M. Vilà, C. A. Wood, A. Zenetos, and

W. Rabitsch. 2019. Developing a list of invasive alien species likely to threaten biodiversity and ecosystems in the European Union. Global Change Biology **25**:1032-1048.

Russell, J. C., S. D. Miller, G. A. Harper, H. E. MacInnes, M. J. Wylie, and R. M. Fewster. 2010. Survivors or reinvaders? Using genetic assignment to identify invasive pests following eradication. Biological Invasions **12**:1747-1757.

Schindler, S., B. Staska, M. Adam, W. Rabitsch, and F. Essl. 2015. Alien species and public health impacts in Europe: a literature review. NeoBiota **27**:1.

Scott, R., A. Zhan, E. A. Brown, F. J. J. Chain, M. E. Cristescu, R. Gras, and H. J. MacIsaac. 2018. Optimization and performance testing of a sequence processing pipeline applied to detection of nonindigenous species. Evolutionary Applications **11**:891-905.

Seebens, H., S. Bacher, T. M. Blackburn, C. Capinha, W. Dawson, S. Dullinger, P. Genovesi, P. E. Hulme, M. van Kleunen, I. Kühn, J. M. Jeschke, B. Lenzner, A. M. Liebhold, Z. Pattison, J. Pergl, P. Pyšek, M. Winter, and F. Essl. 2021. Projecting the continental accumulation of alien species through to 2050. Global Change Biology **27**:970-982.

Seebens, H., T. M. Blackburn, E. E. Dyer, P. Genovesi, P. E. Hulme, J. M. Jeschke, S. Pagad, P. Pyšek, M. Winter, M. Arianoutsou, S. Bacher, B. Blasius, G. Brundu, C. Capinha, L. Celesti-Grapow, W. Dawson, S. Dullinger, N. Fuentes, H. Jäger, J. Kartesz, M. Kenis, H. Kreft, I. Kühn, B. Lenzner, A. Liebhold, A. Mosena, D. Moser, M. Nishino, D. Pearman, J. Pergl, W. Rabitsch, J. Rojas-Sandoval, A. Roques, S. Rorke, S. Rossinelli, H. E. Roy, R. Scalera, S. Schindler, K. Štajerová, B. Tokarska-Guzik, M. van Kleunen, K. Walker, P. Weigelt, T. Yamanaka, and F. Essl. 2017. No saturation in the accumulation of alien species worldwide. Nature Communications **8**:14435.

Shackleton, R. T., T. Adriaens, G. Brundu, K. Dehnen-Schmutz, R. A. Estévez, J. Fried, B. M. H. Larson, S. Liu, E. Marchante, H. Marchante, M. C. Moshobane, A. Novoa, M. Reed, and D. M. Richardson. 2019. Stakeholder engagement in the study and management of invasive alien species. Journal of Environmental Management **229**:88-101.

Sigsgaard, E. E., M. R. Jensen, I. E. Winkelmann, P. R. Møller, M. M. Hansen, and P. F. Thomsen. 2020. Population-level inferences from environmental DNA—Current status and future perspectives. Evolutionary Applications **13**:245-262.

Sigsgaard, E. E., I. B. Nielsen, S. S. Bach, E. D. Lorenzen, D. P. Robinson, S. W. Knudsen, M. W. Pedersen, M. Jaidah, L. Orlando, E. Willerslev, P. R. Moller, and P. F. Thomsen. 2017. Population characteristics of a large whale shark aggregation inferred from seawater environmental DNA. Nature Ecology & Evolution **1**.

Simberloff, D. 2020. Maintenance management and eradication of established aquatic invaders. Hydrobiologia.

Sorte, C. J. B., S. L. Williams, and J. T. Carlton. 2010. Marine range shifts and species introductions: comparative spread rates and community impacts. Global Ecology and Biogeography **19**:303-316.

Spens, J., A. R. Evans, D. Halfmaerten, S. W. Knudsen, M. E. Sengupta, S. S. T. Mak, E. E. Sigsgaard, M. Hellström, and D. Yu. 2017. Comparison of capture and storage methods for aqueous macrobial eDNA using an optimized extraction protocol: advantage of enclosed filter. Methods in Ecology and Evolution **8**:635-645.

Steinke, D., S. W. J. Prosser, and P. D. N. Hebert. 2016. DNA Barcoding of Marine Metazoans. Pages 155-168 *in* S. J. Bourlat, editor. Marine Genomics: Methods and Protocols. Springer New York, New York, NY.

Sunday, J. M., A. E. Bates, and N. K. Dulvy. 2012. Thermal tolerance and the global redistribution of animals. Nature Climate Change **2**:686-690.

Sunday, J. M., G. T. Pecl, S. Frusher, A. J. Hobday, N. Hill, N. J. Holbrook, G. J. Edgar, R. Stuart-Smith, N. Barrett, T. Wernberg, R. A. Watson, D. A. Smale, E. A. Fulton, D. Slawinski, M. Feng, B. T. Radford, P. A. Thompson, and A. E. Bates. 2015. Species traits and climate velocity explain geographic range shifts in an ocean-warming hotspot. Ecology Letters **18**:944-953.

Svane, I., and C. M. Young. 1989. The ecology and behaviour of ascidian larvae. Oceanography and Marine Biology: an Annual Review **27**:45-90.

Tarjuelo, I., D. Posada, K. Crandall, M. Pascual, and X. Turon. 2001. Cryptic species of Clavelina (*Ascidiacea*) in two different habitats: harbours and rocky littoral zones in the northwestern Mediterranean. Marine Biology **139**:455-462.

Thresher, R. E., and A. M. Kuris. 2004. Options for Managing Invasive Marine Species. Biological Invasions **6**:295-300.

Tsuji, S., M. Miya, M. Ushio, H. Sato, T. Minamoto, and H. Yamanaka. 2020. Evaluating intraspecific genetic diversity using environmental DNA and denoising approach: A case study using tank water. Environmental DNA **2**:42-52.

Turon, X., A. Antich, C. Palacín, K. Præbel, and O. S. Wangensteen. 2020. From metabarcoding to metaphylogeography: separating the wheat from the chaff. Ecological Applications **30**:e02036.

Uchii, K., H. Doi, and T. Minamoto. 2016. A novel environmental DNA approach to quantify the cryptic invasion of non-native genotypes. Molecular Ecology Resourses **16**:415-422.

Vilà, M., C. Basnou, P. Pyšek, M. Josefsson, P. Genovesi, S. Gollasch, W. Nentwig, S. Olenin, A. Roques, D. Roy, P. E. Hulme, and D. partners. 2010. How well do we understand the impacts of alien species on ecosystem services? A pan-European, cross-taxa assessment. Frontiers in Ecology and the Environment **8**:135-144.

Weitemier, K., B. E. Penaluna, L. L. Hauck, L. J. Longway, T. Garcia, and R. Cronn. 2021. Estimating the genetic diversity of Pacific salmon and trout using multigene eDNA metabarcoding. Molecular Ecology **In Press**.

Williams, S. L., I. C. Davidson, J. R. Pasari, G. V. Ashton, J. T. Carlton, R. E. Crafton, R. E. Fontana, E. D. Grosholz, A. W. Miller, G. M. Ruiz, and C. J. Zabin. 2013. Managing Multiple Vectors for Marine Invasions in an Increasingly Connected World. BioScience **63**:952-966.

Zhan, A., E. Briski, D. G. Bock, S. Ghabooli, and H. J. MacIsaac. 2015. Ascidians as models for studying invasion success. Marine Biology **162**:2449-2470.

Zhan, A., M. Hulák, F. Sylvester, X. Huang, A. A. Adebayo, C. L. Abbott, S. J. Adamowicz, D. D. Heath, M. E. Cristescu, and H. J. MacIsaac. 2013. High sensitivity of 454 pyrosequencing for detection of rare species in aquatic communities. Methods in Ecology and Evolution **4**:558-565.

4.8 Supplementary Information

4.8.1 Supplementary Table 1

Sampling sites, latitude, longitude and date of sampling for non-native ascidian species at South African sites used in biodiversity and genetic surveys.

Site Name	Site Code	Latitude (S)	Longitude (E)	Sampling Date
Saldanha Bay	SB	-33.018908	17.950512	14-Oct-17
Club Mykonos Marina	SM	-33.046315	18.04028	15-Oct-17
V&A Waterfront Marina	TB	-33.909055	18.419893	18-Oct-17
Hout Bay Yacht Club	HB	-34.049767	18.348042	19-Oct-17
Mossel Bay Marina	MB	-34.178386	22.144918	22-Oct-17
Knysna	KN	-34.041086	23.043592	24-Oct-17
Port Elizabeth Marina	PE	-33.96692	25.634461	27-Oct-17
Bushmans River (Kenton Marina)	BR	-33.679558	26.655841	30-Oct-17
Port Alfred Marina	PA	-33.593546	26.892084	31-Oct-17
East London Marina	EL	-33.024163	27.896269	02-Nov-17
Durban Marina	DU	-29.862663	31.021903	08-Nov-17
Richards Bay Marina	RB	-28.793908	32.079184	10-Nov-17

4.8.2 Supplementary Table 2

__a__ PCR conditions for each species with time shown in minutes and temperature in degrees Celsius. __b__ Sources and sequences (5' to 3' orientation) of primers used in cytochrome c oxidase subunit I sequencing of South African non-native ascidian species, further publication details for primers are provided in the main manuscript references, digital object identifiers are provided in the table.

a

Species	***Denaturation***		***Cycles***							***Extension***	
	Temp	**Time**	**Temp**	**Time**	**Temp**	**Time**	**Temp**	**Time**	**Cycles**	**Temp**	**Time**
Ciona intestinalis	95	10:00	95	0:50	48	0:30	72	1:00	37	72	7:00
Clavelina lepidaformis	95	10:00	95	0:30	55	0:30	72	1:00	37	72	7:00
Microcosmus Squamiger	95	10:00	95	0:30	50	0:30	72	1:00	35	72	7:00
Styela plicata	95	10:00	95	0:30	50	0:30	72	1:00	35	72	7:00

b

Species	*PrimerF*	*Forward Sequence 5' – 3'*	*PrimerR*	*Reverse Sequence 5' – 3'*	*Source*	*DOI*
Ciona intestinalis	NydamHarrisF	GAGTAAGAACTGGRTGRACAGTTTAYCCTCC	NydamHarrisR	ATTAAAACTTAATCTAGTAAAAAGAGGRRATCAATGG	Nydam & Harrison, 2010	10.1007/s00227-007-0617-0
Clavelina lepadiformis	Clav2001F	GTACTGAGCTTTCACAAACTGGGCAAT	Clav2001R	TGAAAAAGAATAGGATCTCTCCTTCC	Tarjuelo et al. 2001	10.1007/s002270100587
Microcosmus Squamiger	Mc_F	CCTCTGGGTGCCTAAAAAATCA	Mc_R	AGATGTGAATTAAGCGTTGGC	This work	-
Styela plicata	TunF1	TDTCAACDAATCATAARGATATTRG	TunR1	TAAACYTCAGGATGTCYAAARAAYCA	Steinke et al.	10.1007/978-1-4939-3774-5_10

4.8.3 Supplementary Table 3

Analysis of molecular variance (AMOVA) partitioning variance found in COI sequences by sampling occasion (year) site and within samples. Model outputs are shown for all species.

Ciona robusta	**Df**	**SumSq**	**Mean Sq**	**P value**	**% variance**
Between sampling years	1	29.40128	29.401285	0.20879121	5.879724
Between sites within year	9	135.00643	15.000715	0.000999	27.8254
Within samples	264	348.71228	1.32088	0.000999	66.294876
Total	274	513.12	1.872701		100

Clavelina lepadiformis	**Df**	**SumSq**	**Mean Sq**	**P value**	**% variance**
Between sampling years	1	6.30415	6.3041504	0.18481519	1.489209
Between sites within year	17	62.33149	3.6665583	0.000999	19.434993
Within samples	452	235.94185	0.5219953	0.000999	79.075799
Total	470	304.57749	0.6480372		100

Styela plicata	**Df**	**SumSq**	**Mean Sq**	**P value**	**% variance**
Between sampling years	1	15.92812	15.928116	0.6953047	-8.706159
Between sites within year	7	674.5894	96.369914	0.000999	42.221948
Within samples	219	1244.70179	5.68357	0.000999	66.484211
Total	227	1935.2193	8.525195		100

Microcosmus squamiger	**Df**	**SumSq**	**Mean Sq**	**P value**	**% variance**
Between sampling years	1	0.867717	0.867717	0.98201798	-1.065651
Between sites within year	10	29.099302	2.90993	0.02697303	3.166193
Within samples	234	413.081762	1.765307	0.04295704	97.899458
Total	245	443.04878	1.808362		100

4.8.4 Supplementary Note 1

Comparisons between rapid assessment surveys and eDNA metabarcoding derived relative abundance measures for ascidians across the coast of South Africa.

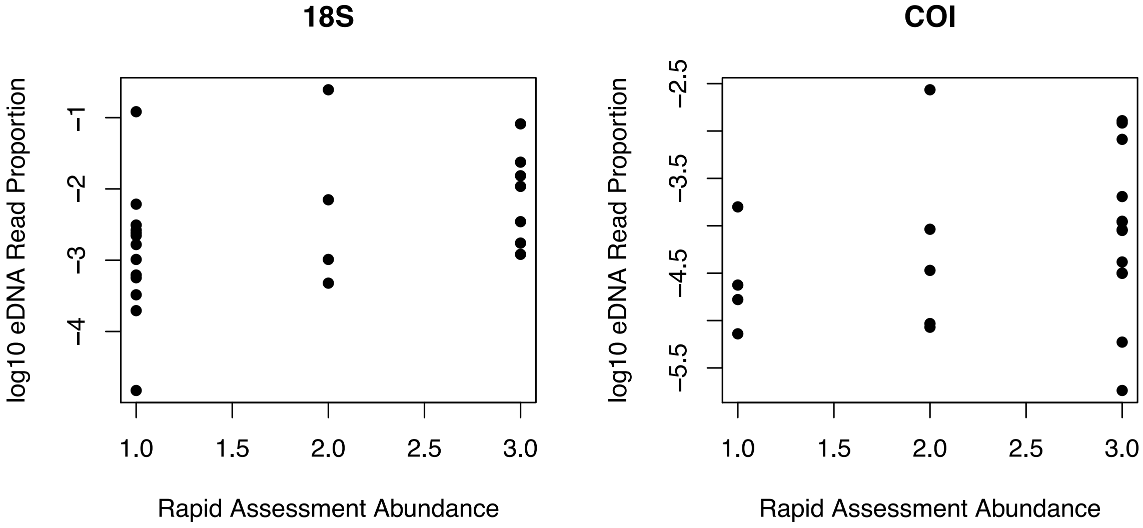
In order to explore if there is a relationship between the abundance assessments from the rapid assessment surveys and the eDNA metabarcoding surveys the 18S & COI metabarcoding data was parsed as in the main manuscript for the nine non-native ascidians found across the South Africa coastline [following (Rius et al. 2014)]. These species are *Ascidia sydneiensis, Ascidiella aspersa, Asterocarpa humilis, Botryllus schlosseri, Ciona robusta, Clavelina lepadiformis, Diplosoma listerianum, Microcosmus squamiger, Styela plicata*. For the 18S data sequences could be assigned to *A. aspersa, A. humulis, B. schlosseri, M. squamiger, S. plicata*, for the COI data sequences could be assigned to all species except *A. sydneiensis* and *A. humilis*. All zero abundance measures were removed from the analysis. Linear models of log10 transformed read proportions against categorical rapid assessment for all species combined were used to assess the relationship between metabarcoding reads and rapid assessment abundance (see Supplementary Note 1 Table 1 below). No significant relationship was found between metabarcoding reads and rapid assessment abundance in either the 18S or COI data ($p > 0.05$).

Supplementary Note 1 Table 1 *Table detailing model outputs for ordinary least squares regression models between rapid assessment survey abundance and log10 eDNA metabarcoding relative read abundance. Data is shown for both 18S and COI.*

18S	Estimate	Std. Error	t value	p value
(Intercept)	-3.2924	0.399	-8.252	3.51E-08
Rapid Assessment	0.4187	0.2038	2.055	0.052
F value	4.221	1 and 22 Df		

COI	Estimate	Std. Error	t value	p value
(Intercept)	-4.7869	0.5733	-8.35	5.98E-08
Rapid Assessment	0.2425	0.2264	1.071	0.297
F value	1.147	1 and 20 Df		

Supplementary Note 1 Figure 1 *Plots showing rapid assessment survey abundance against log10 eDNA metabarcoding relative read abundance for 18S (left) and COI (right). Rapid assessment abundance is measured in categories of 1(scarce), 2(common) or 3(dominant) following the main text.*

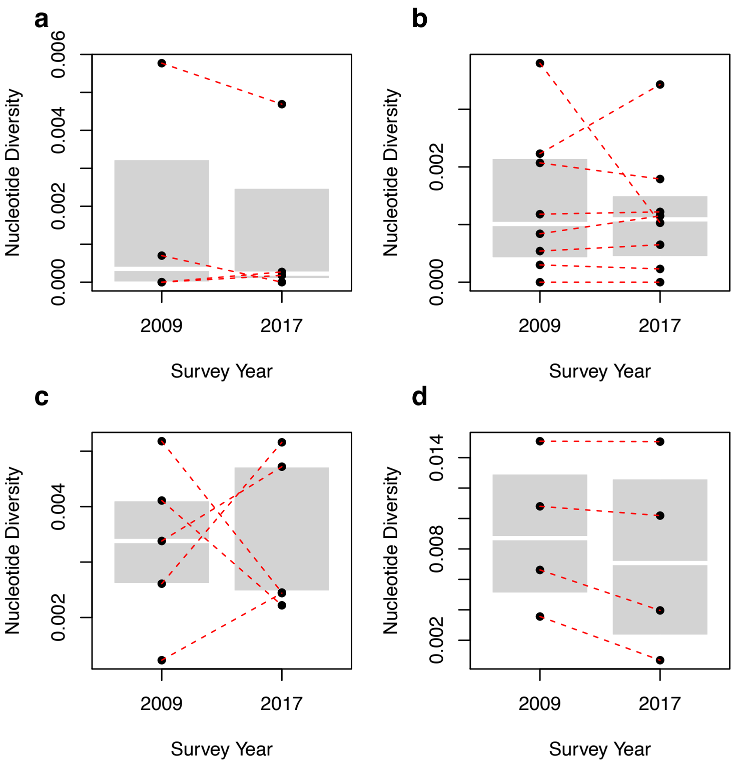


4.8.5 Supplementary Note 2

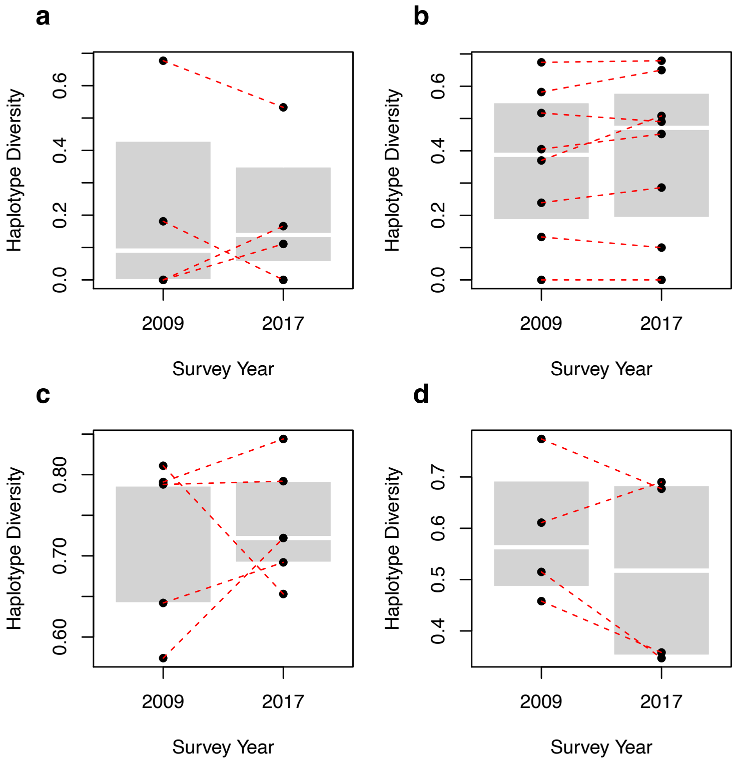
Analyses of nucleotide and haplotype diversity for non-native ascidians across South Africa.

In order to explore if there is a significant change in genetic diversity between sampling occasions data was subset to include only sites were genetic data was collected during both sampling occasions (2009 & 2017) for all species. Nucleotide and haplotypic diversity were calculated using the *nuc.div* and *hap.div* functions from the *pegas* package. Paired samples T-tests were then performed on the diversity estimates to evaluate if there was a significant difference in both haplotype and nucleotide diversity between sampling occasions across the different species; *Ciona robusta; Clavelina lepadiformis; Microcosmus squamiger; Styela plicata.*

Supplementary Note 2 Figure 1 *Nucleotide diversity for 2009 and 2017 sampling occasions from sequenced tissue samples. Observations form the same site are connected with dashed red lines, the grey boxes indicate 25% and 7% quartiles with the median value shown by a white dash.* ***a*** *Ciona robusta* ***b*** *Clavelina lepadiformis* ***c*** *Microcosmus squamiger* ***d*** *Styela plicata*



Supplementary Note 2 Figure 2 *Haplotype diversity for 2009 and 2017 sampling occasions from sequenced tissue samples. Observations from the same site are connected with dashed red lines, the grey boxes indicate 25% and 7% quartiles with the median value shown by a white dash.* ***a*** *Ciona robusta* ***b*** *Clavelina lepadiformis* ***c*** *Microcosmus squamiger* ***d*** *Styela plicata*



Supplementary Note 2 Table 1 *Model output from paired sample T-test conducted on haplotype and nucleotide diversity between sampling occasions. Each row indicates a different model with species and statistic note for each case. Model t value, degrees of freedom (Df) and p values are shown.*

Species	**Statistic**	**t value**	**Df**	**p value**
Ciona robusta	Haplotype diversity	0.13646	3	0.9001
Ciona robusta	Nucleotide diversity	1.0027	3	0.3899
Clavelina lepadiformis	Haplotype diversity	-1.5304	7	0.1698
Clavelina lepadiformis	Nucleotide diversity	0.45424	7	0.6634
Microcosmus squamiger	Haplotype diversity	-0.3868	4	0.7186
Microcosmus squamiger	Nucleotide diversity	-0.094269	4	0.9294
Styela plicata	Haplotype diversity	1.3548	3	0.2685
Styela plicata	Nucleotide diversity	2.1515	3	0.1205

4.8.6 Supplementary Note 3

Evaluation of a raw read mapping approach for the detection of non-native ascidians across South Africa

Previous work has identified that bioinformatic methods and parameters used in the analysis of DNA metabarcoding data can have an effect on biodiversity metrics (Scott et al. 2018, Antich et al. 2021). Therefore, a raw read mapping approach was tested as follows. For *Ciona robusta*, *Clavelina lepadiformis, Styela plicata* and *Microcosmus squamiger* all the available cytochrome c oxidase subunit I gene (COI) sequences on NCBI Nucleotide database were downloaded on 11/05/2021. Raw eDNA metabarcoding reads from the COI gene region sequenced across the sites in the main manuscript for each sample were then mapped against the downloaded COI sequences using the --usearch_global setting of vsearch (v2.10.4) with the following parameters *-id 0.97 -strand both -maxhits 1*. The resultant mappings were then summed for each species in each site to produce Table 1 below.

	SB	SM	TB	HB	MB	KN	PE	BR	PA	EL	DU	RB
Ciona robusta	1	0	1	1*	0	1*	0	0	0	0	0	0
Clavelina lepadiformis	1	1	1	1*	1	1	1*	0	1	0	0	0
Microcosmus squamiger	0	0	0	0	1	1	1	1*	1	1	0	1
Styela plicata	0	1	1!*	1!*	1	1	1	0	0	1!*	1!*	1

Supplementary Note 3 Table 1

Table showing presence (1) and absence (0) of ascidians in sites across South Africa from eDNA metabarcoding data analysed by mapping to COI regions. Underlined values show detections previously not found when analysing the dataset using a clustering and QC pipeline. Values marked with '!' showed read counts lower than those found in negative control samples. Values marked with '' correspond with singleton detections that are typically discarded in bioinformatic pipelines. Site abbreviations as in Supplementary Table 1.*

The results show several instances of positive detections not found when using a clustering/denoising approach that match up with positive detections from rapid assessment surveys. However, many of these positive detections correspond with one or two reads, which are typically discarded regardless of bioinformatic pipeline. Additionally, one read from a single negative control sample was mapped to *Styela plicata* which suggests extremely low, but non-zero cross-contamination at some stage across the workflow, which is typical of metabarcoding datasets. This shows that while it may be possible to generate more sensitive incidence data by using a raw mapping approach it may not be advisable as the possibility for false positive errors increases.

4.8.7 Supplementary References

Antich, A., C. Palacin, O. S. Wangensteen, and X. Turon. 2021. To denoise or to cluster, that is not the question: optimizing pipelines for COI metabarcoding and metaphylogeography. BMC Bioinformatics **22**:177.

Rius, M., S. Clusella-Trullas, C. D. McQuaid, R. A. Navarro, C. L. Griffiths, C. A. Matthee, S. von der Heyden, and X. Turon. 2014. Range expansions across ecoregions: interactions of climate change, physiology and genetic diversity. Global Ecology and Biogeography **23**:76-88.

Scott, R., A. Zhan, E. A. Brown, F. J. J. Chain, M. E. Cristescu, R. Gras, and H. J. MacIsaac. 2018. Optimization and performance testing of a sequence processing pipeline applied to detection of nonindigenous species. Evolutionary Applications **11**:891-905.

Chapter 5 Demonstration of the use of environmental DNA for the non-invasive genotyping of a bivalve mollusk, the European Flat Oyster (*Ostrea edulis*)

5.1 Citation

Here I present a full typeset article as published in *Frontiers in Genetics*, the citation is provided below.

Holman LE, Hollenbeck CM, Ashton TJ, Johnston IA (2019*)* Demonstration of the Use of Environmental DNA for the Non-Invasive Genotyping of a Bivalve Mollusk, the European Flat Oyster (Ostrea edulis). *Frontiers in Genetics*. 10:1159.
DOI:10.3389/fgene.2019.01159

frontiers
in Genetics

METHODS
published: 19 November 2019
doi: 10.3389/fgene.2019.01159

Demonstration of the Use of Environmental DNA for the Non-Invasive Genotyping of a Bivalve Mollusk, the European Flat Oyster (*Ostrea edulis*)

Luke E. Holman*[1]*, *Christopher M. Hollenbeck*[2], *Thomas J. Ashton*[2] *and Ian A. Johnston*[2,3]*

[1] School of Ocean and Earth Science, National Oceanography Centre Southampton, University of Southampton, Southampton, United Kingdom, [2] Xelect Ltd, Horizon House, Scotland, United Kingdom, [3] Scottish Oceans Institute, School of Biology, University of St Andrews, Scotland, United Kingdom

OPEN ACCESS

Edited by:
Jesús Fernández,
Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Spain

Reviewed by:
Pierre Boudry,
Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER), France
Juan Andres Rubiolo,
University of Santiago de Compostela, Spain

****Correspondence:***
Luke E. Holman
leh1n15@soton.ac.uk
Ian A. Johnston
ian.johnston@xelect.co.uk

Specialty section:
This article was submitted to Livestock Genomics, a section of the journal Frontiers in Genetics

Received: *04 July 2019*
Accepted: *22 October 2019*
Published: *19 November 2019*

Citation:
Holman LE, Hollenbeck CM, Ashton TJ and Johnston IA (2019) Demonstration of the Use of Environmental DNA for the Non-Invasive Genotyping of a Bivalve Mollusk, the European Flat Oyster (Ostrea edulis). Front. Genet. 10:1159. doi: 10.3389/fgene.2019.01159

Accurate SNP (single nucleotide polymorphism) genotype information is critical for a wide range of selective breeding applications in aquaculture, including parentage assignment, marker-assisted, and genomic selection. However, the sampling of tissue for genetic analysis can be invasive for juvenile animals or taxa where sampling tissue is difficult or may cause mortality (e.g. bivalve mollusks). Here, we demonstrate a novel, non-invasive technique for sampling DNA based on the collection of environmental DNA using European Flat Oysters (*Ostrea edulis)* as an example. The live animals are placed in individual containers until sufficient genetic material is released into the seawater which is then recovered by filtration. We compared the results of tissue and eDNA derived SNP genotype calls using a PCR based genotyping platform. We found that 100% accurate genotype calls from eDNA are possible, but depend on appropriate filtration and the dilution of the sample throughout the workflow. We also developed an additional low-cost DNA extraction technique which provided >99% correct SNP genotype calls in comparison to tissue. It was concluded that eDNA sampling can be used in hatchery and selective breeding programs applicable to any aquatic organism for which direct tissue sampling may result in animal welfare concerns or mortality.

Keywords: broodstock, hatchery management, single nucleotide polymorphism genotyping, mollusk aquaculture, minimally invasive sampling, non-invasive genetic sampling

INTRODUCTION

Molluscan shellfish, such as clams, oysters, mussels, and scallops, represent around 20% of worldwide aquaculture production (FAO, 2019). Although the life cycle has been closed for many mollusks, most aquaculture production is still dependent on unpredictable collection of spat (settled larvae) from the wild. The availability of wild spat can be negatively impacted by overfishing, environmental or trophic changes (Waldbusser et al., 2015; Lagarde et al., 2018), and disease outbreaks (Boudry et al., 1998; Garcia et al., 2011; Murray et al., 2012), none of which are under the control of the producer. There is therefore a trend towards hatchery-based production of juveniles for on-growing in the sea. Hatchery-based production allows for genetic improvement of stock *via* selective breeding, which has the potential to improve economically important traits such as growth and disease resistance by 10%–15% per generation (Hollenbeck and Johnston, 2018).

One particular challenge of molluscan aquaculture is the availability of non-invasive DNA sampling techniques for parentage assignment and advanced marker-assisted or genomic selection strategies. Current DNA collection strategies for mollusks involve the use of anesthetic chemicals to relax internal muscles which opens the shell to enable clipping of internal tissues such as the gill or mantle which are not accessible when the shell is closed (Suquet et al., 2009; Suquet et al., 2010; Mao et al., 2013). An alternative method is to sample internal fluids using a syringe with or without anesthetic (Jones et al., 1993; Kurita and Kijima, 2019). These methods can result in a physiological stress response (Butt et al., 2008; Granados-Amores et al., 2017) and in some cases cause mortality of valuable broodstock (Henley et al., 2006). Furthermore, clipping of internal tissues such as mantle tissue which is rich in mechanoreceptors and chemoreceptors is problematic from an animal welfare perspective, particularly for small species or for individuals that are immunocompromised or in poor condition.

Recent advances in the isolation of environmental DNA (eDNA) potentially offer a non-invasive alternative to tissue sampling (Carroll et al., 2018). eDNA is a polydisperse mixture of nucleic acid containing material shed from an organism and isolated from environmental samples such as sediment or water (Thomsen and Willerslev, 2015; Deiner et al., 2017). The majority of eDNA studies have been used to test ecological hypotheses either by recording the incidence of a single aquatic species using species-specific primers (Collins et al., 2018; Seymour et al., 2018) or many species simultaneously with metabarcoding (Stat et al., 2017; Deiner et al., 2018; Holman et al., 2019). Overall, eDNA has been shown to be highly accurate and at least as sensitive as other biodiversity monitoring techniques (Deiner et al., 2017). Additionally, studies have shown that eDNA can provide population genetic inference both in the laboratory (Espinoza et al., 2017) and in coastal ecosystems (Sigsgaard et al., 2017; Stat et al., 2017). In aquaculture species, eDNA has recently been used for the detection of bacterial and parasitic diseases (Nguyen et al., 2018; Peters et al., 2018).

The aim of the present study was to determine whether eDNA could be used to genotype individual bivalves at multiple SNP loci with the accuracy required for parentage assignment. We tested both low cost and archive grade eDNA extraction methods and developed a protocol that achieved 100% accurate genotype calls in comparison to tissue samples from the same individuals. The use of eDNA for the non-invasive genotyping of bivalve broodstock and their offspring represents an important new tool for the development of hatchery-based selective breeding programs.

MATERIALS AND METHODS

Animals and Water Sampling

Six European flat oysters (*Ostrea edulis* L.), 30–80 g, were acclimatized in a 50 L seawater aquarium at 16°C for 60 days, with 700 L/hour external filtration (Ehiem, Deizisau, Germany). Twenty percent of aquarium water was replaced weekly with fresh sea water. The oysters were from aquaculture populations obtained from Loch Nell Oysters (Argyll, UK) and are derived from native stock from the Argyll area. During acclimation, the oysters were fed a maintenance diet of powdered algal biomass (Megatech Research, Switzerland). Each oyster was externally rinsed with reverse osmosis (RO) filtered water, then placed into a polypropylene vessel with 500 ml seawater made from artificial salt (Red Sea Aquatics Ltd, London, UK) dissolved in (RO) water to 33ppt salinity. Duplicate water samples of 75 ml were taken from each vessel 72 h after the oyster was introduced. The oysters were sacrificed and a 5 mm^2 section of mantle was dissected and stored in 100% ethanol until DNA extraction. A 75 ml artificial seawater control sample was taken before filling the vessels. All 75 ml water samples (experimental and control) were filtered using a vacuum filtration manifold and 47 mm 0.45 µm Cellulose Nitrate filters (Sartorius, Göttingen, Germany) in a glass housing.

All reused equipment was soaked in 0.5% sodium hypochlorite solution (household bleach solution diluted 1 in 10 with RO water) for 1 hour before the start of the experiment. Filtration equipment was thoroughly washed between sampling and 100 ml bleach solution (as above) followed by 200 ml RO water was filtered between every sample to minimize the possibility of cross contamination.

DNA Extraction

Approximately 25 mg mantle tissue was dissected and finely sliced with a sterile scalpel. DNA was extracted using the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) under the manufacturer's recommended protocol with a final DNA elution in 100 µl of PCR cert. water. All eDNA filters were sliced into ~3 mm sections using a sterile scalpel. One replicate at each sample point was subject to DNA extraction using the DNeasy Blood and Tissue kit. Briefly, 80 µl of Proteinase K solution (20 mg/ml) and 720 µl of Qiagen ATL Buffer was added to each sliced filter and thoroughly vortexed followed by overnight digestion at 56°C. Five hundred microliters of lysate was mixed with 500 µl of Buffer AL and 500 µl of 100% ethanol, DNA extraction proceeded as in the manufacturer's protocol, with the entire 1,500 µl of lysate being passed through the extraction column. DNA was eluted in 60 µl of PCR grade water. A second eDNA replicate was subject to a crude, low-cost DNA extraction (Walsh et al., 1991) in which 800 µl of 10% Chelex 100 (Sigma-Aldrich, St Louis, USA) suspension containing 0.2 mg/ml Proteinase K was added to the filter. The mixture was then thoroughly vortexed and incubated at 56°C for 60 minutes, 60 µl of lysate was removed and incubated at 95°C for 10 min and then stored at -20°C. DNA concentration was calculated using the Quant-iT PicoGreen dsDNA Assay Kit (ThermoFisher Scientific, Waltham, USA) on the Agilent (Santa Clara, USA) Mx3000P qPCR instrument using the protocol described in Blotta et al. (2005).

Genotyping

O. edulis SNP sequence data from Gutierrez et al. (2017) was sorted by mean minor allele frequency (MAF) across the discovery populations. SNPs used in parentage assignment are frequently selected for high MAF i.e. > 0.2 for better discrimination between individuals (Holman et al., 2017). This property also makes them appropriate for checking genotype accuracy. Therefore the 16 SNPs with the highest mean MAF were sent to the Fluidigm D3 assay design portal for synthesis of Fluidigm SNP Type genotyping assays (**Supplementary Table 1**). Genotyping proceeded with the Fluidigm EP1 platform using the manufacturer's protocols (Fluidigm Ltd., San Francisco, USA). Briefly, template was subject to a multiplex PCR containing primers for all 16 target regions. This Specific Target Amplification step (STA) increases target copy number, making the template amenable to amplification in microfluidic chambers. The STA product was then loaded on to a Fluidigm 96.96 Dynamic Array genotyping chip along with assays in sextuplicate, followed by a PCR at manufacturer recommended conditions and imaging on the Fluidigm EP1 data collection system. The quantity of target DNA in environmental DNA is highly dilute compared to DNA template samples typically used on genotyping platforms (e.g. tissue, cell culture). Therefore, sample dilutions recommended in the manufacturer's protocol were predicted to dramatically alter the accuracy of genotype calls. Dilution of DNA samples before STA is important to minimize the amount of potentially PCR inhibiting co-purified contaminants from DNA extractions, while also transferring an appropriate amount of DNA template for the PCR reaction. Dilution after the STA is a balance between diluting unused reagents from the PCR while transferring enough target copies for successful fluorescence-based genotyping. With these limitations in mind, several different dilutions with RO water were trialled as shown in **Figure 1**.

SNP Genotyping Analysis

SNP genotypes were called using k-means clustering under the default settings in the Fluidigm Genotyping Analysis Software. Calls were checked manually to ensure clustering was performed appropriately within dilutions. Genotypes for eDNA samples were called in relation to the tissue samples from the equivalent dilution.

To evaluate eDNA genotyping success, four metrics were evaluated. Overall SNP call rate was calculated as the proportion of allele calls across the 96 assays (16 SNPs, 6 replicates per SNP) for the eDNA sample that matched the tissue sample. Replicate SNP reactions were considered independent and "no calls" were included in calculations. This metric evaluated both the overall genotyping success of the Fluidigm EP1 system and also the congruence between results from eDNA and tissue DNA samples.

SNP genotype success was calculated by randomly sampling data points from the six replicates to simulate different levels of replication. Scenarios with three, two and a single replicate were simulated one hundred times. In each simulation the majority genotype, excluding "no calls", from the subsample was compared to the tissue derived genotype. In the case when there were two or more conflicting genotypes of equal frequency in a simulation, the call was marked incorrect.

RESULTS

No mortality was recorded and oysters showed no visible sign of spawning before, during, or after eDNA sampling. DNA was successfully extracted from all samples. Oyster tissue samples assayed with Picogreen contained 45.9 ± 21.3 (s.d.) ng/µl of dsDNA, control samples contained 0.48 ± 0.47 ng/µl of dsDNA, and eDNA samples from the DNeasy extraction contained 115.4 ± 50.6 ng/µl of dsDNA.

Out of the 16 trialed assays two failed to produce any identifiable clusters, indicating no polymorphism among the tested oysters or a non-functioning assay. In all cases the tissue samples for the remaining 14 assays gave high quality clusters reliably identified using k-means clustering (SNP call data provided in **Supplementary Table 2**).

Across all time points and samples, the average correct genotyping reaction call rate for eDNA samples compared to the tissue samples was 78.1 ± 24.8% (s.d.) with a maximum correct call rate of 100.0% and a minimum correct call rate of 19.0% (**Supplementary Figure 1**).

Across the 100 simulated genotyping scenarios, dilutions D2 and C1 provided the highest accuracy between eDNA and tissue DNA genotypes for the DNeasy and Chelex extractions, respectively, as shown in **Figure 1**. For the D2 dilution, all 100 simulated scenarios under duplicate and triplicate replication gave 100% correct genotypes, with an average of 99.4% accuracy across the 100 scenarios with a single replicate (see **Supplementary Figure 2** for duplicate simulation results). The C1 dilution gave a mean accuracy of 95.8%, 98.5%, and 99.2% across the 100 scenarios with one, two, and three replicates, respectively. As shown in **Figure 1**, all other Chelex dilutions provided poor accuracy across the scenarios with mean values of less than 90% in all cases.

DISCUSSION

Here we show that the collection of eDNA can be used to accurately genotype bivalve mollusks and potentially other aquatic organisms. The influence of DNA extract dilution on genotyping accuracy was assessed to produce a practical protocol for the European flat oyster that can be used by researchers and aquaculture professionals as a template to develop viable alternatives to invasive tissue sampling in similar species. We also demonstrated that eDNA extracted using this protocol is of sufficient quality and quantity for multi-locus genotyping, which is necessary for most applications in aquaculture breeding programs.

The protection offered by an external shell structure provides an evolutionary advantage for many invertebrate taxa. However, in the context of selective breeding programs this is a disadvantage, increasing the force or invasiveness required to sample DNA compared to organisms with no shell. Invasive methods involving the removal of internal tissue or fluid are routinely reported in marine and freshwater mussels (Yanick and Heath, 2000; Gustafson et al., 2005; Henley et al., 2006), oysters (Culloty and Mulcahy, 1992; Suquet et al., 2009; Suquet et al., 2010; Lokmer and Wegner, 2015), and scallops (Mao et al., 2013). Although, methods have been developed for minimally invasive DNA sampling of finfish using external mucus (Le Vin et al., 2011; Taslima et al., 2016), non-invasive methods are lacking for mollusks. The method presented here involves no contact beyond the handling required to place and remove the individual from the sterile seawater chamber, potentially saving on labor

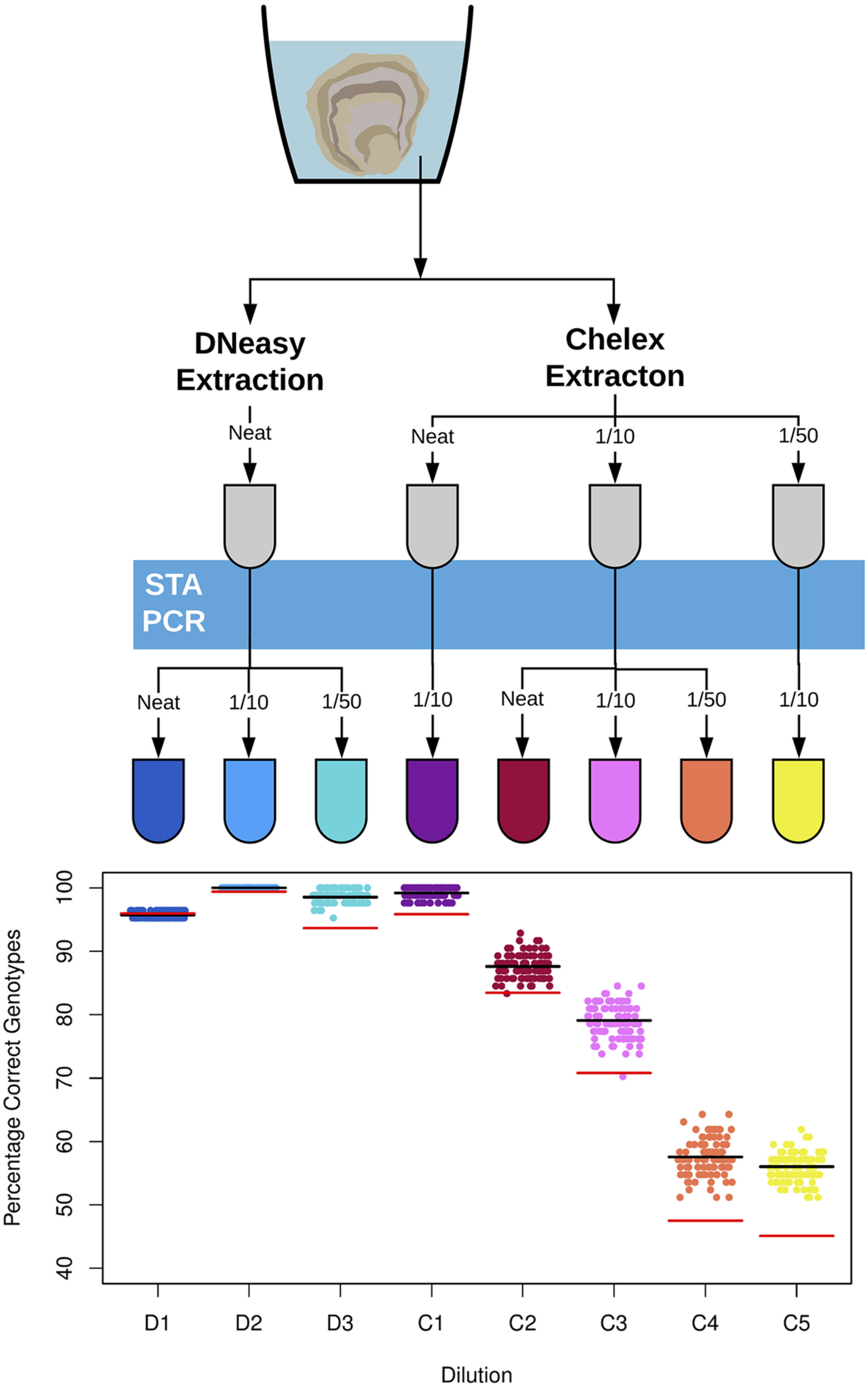


FIGURE 1 | Diagram (top) detailing dilutions of DNA template from different DNA extraction techniques used for SNP genotyping of environmental DNA for *Ostrea edulis*. The two DNA extraction methods, the Qiagen DNeasy Blood and Tissue Kit and the crude Chelex extraction, are shown at the top. DNA samples are subject to dilution step before and after the Specific Target Amplification PCR (STA PCR), here shown as a blue bar. All dilutions were with PCR grade water and samples that were not diluted are labelled 'neat'. A scatter plot (bottom) shows the percentage of correct environmental DNA derived genotype calls in comparison to the tissue extractions for the eight dilutions. Each point represents the total result derived from a random sample of three replicates per individual oyster genotype. The black line indicates the average percentage correct eDNA genotypes across the 100 random samples. The red line indicates the average percentage correct eDNA genotypes across the 100 samples if only a single replicate is used for each individual oyster genotype.

costs as well as achieving a higher animal welfare standard. The equipment required is readily available, relatively inexpensive and provides high quality DNA for downstream applications.

The eDNA genotyping method presented here was developed using a species of comparatively little aquaculture interest. A question that remains is to what extent this method is applicable to other mollusks and aquaculture species. Previous studies have documented that species relevant to aquaculture such as salmonid fish (Atkinson et al., 2018), oysters (Holman et al., 2019), and scallops (Bayer et al., 2019) all produce a sufficient quantity of eDNA for sensitive detection in ecological experiments. Difficulty has only been documented in isolating sufficient eDNA in an invasive crab species (Forsström and Vasemägi, 2016). We can therefore expect almost all aquaculture species to produce eDNA of sufficient yield for the method presented here to provide a non-invasive DNA sample for downstream genetic inference.

Ecological studies have shown a dramatic effect of DNA extraction technique on the results of both eDNA metabarcoding (Djurhuus et al., 2017; Deiner et al., 2018) and qPCR experiments (Hinlo et al., 2017). Our data corroborate these findings, showing that two different eDNA extraction techniques provide variable success in fluorescence-based genotyping, and that a column-based extraction provides greater average genotype accuracy in comparison to a crude lysis technique, albeit at higher cost. We additionally found that dilution of pre- and post-PCR products had an effect on correct genotype calls, decreasing the mean accuracy of genotype calls by over 40% in Chelex extractions. Studies have shown that dilution of eDNA samples has a negative effect on species detectability (McKee et al., 2015; Piggott, 2016), but little work has explored how sample dilution affects SNP genotype accuracy. These results therefore indicate that eDNA users should be wary of diluting samples for accurate genotyping of SNPs or in the estimation of haplotype frequencies.

Small panels of 100–500 SNP markers are the genetic markers of choice for parentage assignment and the determination of relatedness in modern aquaculture breeding programs, allowing both the estimation of breeding values and the control of inbreeding (Robledo et al., 2018). Large panels of SNPs (1,000–50,000) have been used in aquaculture for generating linkage maps (Li and He, 2014), estimating trait heritability (Gutierrez et al., 2018), quantitative trait locus (QTL) mapping (Jiao et al., 2014), genomic selection (Palaiokostas et al., 2016), and most recently the estimation of effective population size (N_e) (D'Ambrosio et al., 2019). SNP arrays (a florescence-based DNA microarray SNP genotyping method) are commonplace for genotyping thousands of SNPs in advanced aquaculture breeding programs. Further work should investigate if the proportion of total isolated eDNA corresponding with the target organism (and not associated bacteria) is of sufficient quality and quantity for other applications than parentage assignment.

Overall, the use of eDNA for SNP genotyping described here will facilitate broodstock management and animal welfare in delicate or hard to sample animals which are enclosed in an external shell or exoskeleton by reducing handling stress and associated mortality.

DATA AVAILABILITY STATEMENT

Raw Illumina sequencing data used to identify variants in Gutierrez et al. (2017) that are genotyped in this work is available at the European nucleotide archive under accession number PRJEB20253 (http://www.ebi.ac.uk/ena/data/view/PRJEB20253). All other datasets used in the study are contained in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

LH, TA, and IJ contributed to the conception and design of the study. LH performed the laboratory work, analyzed the data, and produced the figures. LH and CH wrote an initial draft manuscript. All authors contributed to manuscript revision, read, and approved the final version.

FUNDING

This project has received funding from the European Marine Biological Research Infrastructure Cluster (EMBRIC) project funded by the European Union's Horizon 2020 research and innovation program under grant agreement No 654008. LH was supported by the Natural Environmental Research Council (grant number NE/L002531/1).

ACKNOWLEDGMENTS

The authors thank Rachael Wilbourn and Francesca Tedeschi for assistance in the laboratory and Alejandro Gutierrez for the provision of contig data used in assay design.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2019.01159/full#supplementary-material

REFERENCES

Atkinson, S., Carlsson, J. E. L., Ball, B., Egan, D., Kelly-Quinn, M., Whelan, K., et al. (2018). A quantitative PCR-based environmental DNA assay for detecting Atlantic salmon (*Salmo salar L.*). *Aquat. Conserv.: Marine Freshwater Ecosyst.* 28, 1238–1243. doi: 10.1002/aqc.2931

Bayer, S. R., Countway, P. D., and Wahle, R. A. (2019). Developing an eDNA toolkit to quantify broadcast spawning events of the sea scallop *Placopecten magellanicus*: moving beyond fertilization assays. *Marine Ecol. Prog. Ser.* 621, 127–141. doi: 10.3354/meps12991

Blotta, I., Prestinaci, F., Mirante, S., and Cantafora, A. (2005). Quantitative assay of total dsDNA with PicoGreen reagent and real-time fluorescent detection. *Ann. Ist Super Sanita* 41, 119–123.

Boudry, P., Barré, M., and Gérard, A. (1998). Genetic improvement and selection in shellfish: a review based on oyster research and production. *Cah. Options Méditerr* 34, 61–75.

Butt, D., O'connor, S. J., Kuchel, R., O'connor, W. A., and Raftos, D. A. (2008). Effects of the muscle relaxant, magnesium chloride, on the Sydney rock oyster (*Saccostrea glomerata*). *Aquacult.* 275, 342–346. doi: 10.1016/j.aquaculture.2007.12.004

Carroll, E. L., Bruford, M. W., Dewoody, J. A., Leroy, G., Strand, A., Waits, L., et al. (2018). Genetic and genomic monitoring with minimally invasive sampling methods. *Evol. Appl.* 11, 1094–1119. doi: 10.1111/eva.12600

Collins, R. A., Wangensteen, O. S., O'gorman, E. J., Mariani, S., Sims, D. W., and Genner, M. J. (2018). Persistence of environmental DNA in marine systems. *Commun. Biol.* 1, 185. doi: 10.1038/s42003-018-0192-6

Culloty, S. C., and Mulcahy, M. F. (1992). An evaluation of anaesthetics for *Ostrea edulis* (L.). *Aquacult.* 107, 249–252. doi: 10.1016/0044-8486(92)90073-T

D'ambrosio, J., Phocas, F., Haffray, P., Bestin, A., Brard-Fudulea, S., Poncet, C., et al. (2019). Genome-wide estimates of genetic diversity, inbreeding and effective size of experimental and commercial rainbow trout lines undergoing selective breeding. *Genet. Selection Evol.* 51, 26. doi: 10.1186/s12711-019-0468-4

Deiner, K., Bik, H. M., Machler, E., Seymour, M., Lacoursiere-Roussel, A., Altermatt, F., et al. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Mol. Ecol.* 26, 5872–5895. doi: 10.1111/mec.14350

Deiner, K., Lopez, J., Bourne, S., Holman, L., Seymour, M., Grey, E. K., et al. (2018). Optimising the detection of marine taxonomic richness using environmental DNA metabarcoding: the effects of filter material, pore size and extraction method. *Metabarcoding Metagenomics* 2, e28963. doi: 10.3897/mbmg.2.28963

Djurhuus, A., Port, J., Closek, C. J., Yamahara, K. M., Romero-Maraccini, O., Walz, K. R., et al. (2017). Evaluation of Filtration and DNA Extraction Methods for Environmental DNA Biodiversity Assessments across Multiple Trophic Levels. *Front. In Marine Sci.* 4, 314. doi: 10.3389/fmars.2017.00314

Espinoza, G. J., Poland, J. M., and Bremer, J. R. A. (2017). Genotyping live fish larvae: Non-lethal and noninvasive DNA isolation from 3-5 day old hatchlings. *Biotechniques* 63, 181–186. doi: 10.2144/000114598

FAO (2019).*Fisheries and aquaculture software. FishStatJ—software for fishery statistical time series*. Rome, Italy, FAO Fisheries and Aquaculture Department Rome. 21 July 2016 ed.

Forsström, T., and Vasemägi, A. (2016). Can environmental DNA (eDNA) be used for detection and monitoring of introduced crab species in the Baltic Sea?. *Marine Pollution Bull.* 109, 350–355. doi: 10.1016/j.marpolbul.2016.05.054

Garcia, C., Thebault, A., Degremont, L., Arzul, I., Miossec, L., Robert, M., et al. (2011). Ostreid herpesvirus 1 detection and relationship with Crassostrea gigas spat mortality in France between 1998 and 2006. *Vet. Res.* 42, 73. doi: 10.1186/1297-9716-42-73

Granados-Amores, A., Campa-Cordova, A. I., Rodriguez-Jaramillo, C., Mazon-Suastegui, J. M., Tovar-Ramirez, D., and Saucedo, P. E. (2017). Physiological response to relaxation as pre-operative therapy to pearl formation in the winged pearl oyster. *Pteria sterna. Aquacult.* 473, 145–152. doi: 10.1016/j.aquaculture.2017.02.002

Gustafson, L. L., Stoskopf, M. K., Bogan, A. E., Showers, W., Kwak, T. J., Hanlon, S., et al. (2005). Evaluation of a nonlethal technique for hemolymph collection in *Elliptio complanata*, a freshwater bivalve (Mollusca: Unionidae). *Dis. Aquat. Organisms* 65, 159–165. doi: 10.3354/dao065159

Gutierrez, A. P., Turner, F., Gharbi, K., Talbot, R., Lowe, N. R., Peñaloza, C., et al. (2017). Development of a medium density combined-species SNP array for pacific and european oysters (*Crassostrea gigas and Ostrea edulis*). *G GenesGenomesGenetics* 7, 2209–2218. doi: 10.1534/g3.117.041780

Gutierrez, A. P., Bean, T. P., Hooper, C., Stenton, C. A., Sanders, M. B., Paley, R. K., et al. (2018). A genome-wide association study for host resistance to ostreid herpesvirus in pacific oysters (*Crassostrea gigas*). *G Genes Genomes Genet.* 8, 1273–1280. doi: 10.1534/g3.118.200113

Henley, W. F., Grobler, P. J., and Neves, R. J. (2006). Non-invasive method to obtain DNA from freshwater mussels (Bivalvia: Unionidae). *J. Shellfish Res.* 25, 975–977. doi: 10.2983/0730-8000(2006)25[975:NMTODF]2.0.CO;2

Hinlo, R., Gleeson, D., Lintermans, M., and Furlan, E. (2017). Methods to maximise recovery of environmental DNA from water samples. *PloS One* 12, e0179251. doi: 10.1371/journal.pone.0179251

Hollenbeck, C. M., and Johnston, I. A. (2018). Genomic Tools and Selective Breeding in Molluscs. *Front. In Genet.* 9, 253. doi: 10.3389/fgene.2018.00253

Holman, L. E., Garcia De La Serrana, D., Onoufriou, A., Hillestad, B., and Johnston, I. A. (2017). A workflow used to design low density SNP panels for parentage assignment and traceability in aquaculture species and its validation in Atlantic salmon. *Aquacult.* 476, 59–64. doi: 10.1016/j.aquaculture.2017.04.001

Holman, L. E., De Bruyn, M., Creer, S., Carvalho, G., Robidart, J., and Rius, M. (2019). Detection of introduced and resident marine species using environmental DNA metabarcoding of sediment and water. *Sci. Rep.* 9, 11559. doi: 10.1038/s41598-019-47899-7

Jiao, W. Q., Fu, X. T., Dou, J. Z., Li, H. D., Su, H. L., Mao, J. X., et al. (2014). High-resolution linkage and quantitative trait locus mapping aided by genome survey sequencing: building up an integrative genomic framework for a bivalve mollusc. *DNA Res.* 21, 85–101. doi: 10.1093/dnares/dst043

Jones, T. O., Bourne, N. F., Bower, S. M., and Iwama, G. K. (1993). Effect of repeated sampling on haemolymph pH, PO2 and haemocyte activity in the Pacific oyster, *Crassostrea gigas* (Thunberg). *J. Exp. Marine Biol. Ecol.* 167, 1–10. doi: 10.1016/0022-0981(93)90179-R

Kurita, Y., and Kijima, A. (2019). A noninvasive method for extracting bivalve DNA from the water-filled mantle cavity. *Hydrobiologia* 829, 237–243. doi: 10.1007/s10750-018-3835-1

Lagarde, F., Richard, M., Bec, B., Roques, C., Mortreux, S., Bernard, I., et al. (2018). Trophic environments influence size at metamorphosis and recruitment performance of Pacific oysters. *Marine Ecol. Prog. Ser.* 602, 135–153. doi: 10.3354/meps12666

Le Vin, A. L., Adam, A., Tedder, A., Arnold, K. E., and Mable, B. K. (2011). Validation of swabs as a non-destructive and relatively non-invasive DNA sampling method in fish. *Mol. Ecol. Resour.* 11, 107–109. doi: 10.1111/j.1755-0998.2010.02909.x

Li, Y. G., and He, M. X. (2014). Genetic mapping and QTL analysis of growth-related traits in *pinctada fucata* using restriction-site associated DNA sequencing. *PloS One* 9, e111707. doi: 10.1371/journal.pone.0111707

Lokmer, A., and Wegner, K. M. (2015). Hemolymph microbiome of Pacific oysters in response to temperature, temperature stress and infection. *Isme J.* 9, 670–682. doi: 10.1038/ismej.2014.160

Mao, J. X., Lv, J., Miao, Y., Sun, C. S., Hu, L. P., Zhang, R., et al. (2013). Development of a Rapid and Efficient Method for Non-Lethal DNA Sampling and Genotyping in Scallops. *PloS One* 8, e68096s. doi: 10.1371/journal.pone.0068096

Mckee, A. M., Spear, S. F., and Pierson, T. W. (2015). The effect of dilution and the use of a post-extraction nucleic acid purification column on the accuracy, precision, and inhibition of environmental DNA samples. *Biol. Conserv.* 183, 70–76. doi: 10.1016/j.biocon.2014.11.031

Murray, A. G., Marcos-Lopez, M., Collet, B., and Munro, L. A. (2012). A review of the risk posed to Scottish mollusc aquaculture from Bonamia, Marteilia and oyster herpesvirus. *Aquacult.* 370, 7–13. doi: 10.1016/j.aquaculture.2012.09.033

Nguyen, P. L., Sudheesh, P. S., Thomas, A. C., Sinnesael, M., Haman, K., and Cain, K. D. (2018). Rapid Detection and Monitoring of Flavobacterium psychrophilum in Water by Using a Handheld, Field-Portable Quantitative PCR System. *J. Aquat. Anim. Health* 30, 302–311. doi: 10.1002/aah.10046

Palaiokostas, C., Ferraresso, S., Franch, R., Houston, R. D., and Bargelloni, L. (2016). Genomic Prediction of Resistance to Pasteurellosis in Gilthead Sea Bream (*Sparus aurata*) Using 2b-RAD Sequencing. *G Genes Genomes Genet.* 6, 3693–3700. doi: 10.1534/g3.116.035220

Peters, L., Spatharis, S., Dario, M. A., Dwyer, T., Roca, I. J. T., Kintner, A., et al. (2018). Environmental DNA: A new low-cost monitoring tool for pathogens in salmonid aquaculture. *Front. In Microbiol.* 9, 3009. doi: 10.3389/fmicb.2018.03009

Piggott, M. P. (2016). Evaluating the effects of laboratory protocols on eDNA detection probability for an endangered freshwater fish. *Ecol. Evol.* 6, 2739–2750. doi: 10.1002/ece3.2083

Robledo, D., Palaiokostas, C., Bargelloni, L., Martinez, P., and Houston, R. (2018). Applications of genotyping by sequencing in aquaculture breeding and genetics. *Rev. In Aquacult.* 10, 670–682. doi: 10.1111/raq.12193

Seymour, M., Durance, I., Cosby, B. J., Ransom-Jones, E., Deiner, K., Ormerod, S. J., et al. (2018). Acidity promotes degradation of multi-species environmental DNA in lotic mesocosms. *Commun. Biol.* 1, 4. doi: 10.1038/s42003-017-0005-3

Sigsgaard, E. E., Nielsen, I. B., Bach, S. S., Lorenzen, E. D., Robinson, D. P., Knudsen, S. W., et al. (2017). Population characteristics of a large whale shark aggregation inferred from seawater environmental DNA. *Nat. Ecol. Evol.* 1, 0004. doi: 10.1038/s41559-016-0004

Stat, M., Huggett, M. J., Bernasconi, R., Dibattista, J. D., Berry, T. E., Newman, S. J., et al. (2017). Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. *Sci. Rep.* 7, 12240. doi: 10.1038/s41598-017-12501-5

Suquet, M., De Kermoysan, G., Araya, R. G., Queau, I., Lebrun, L., Le Souchu, P., et al. (2009). Anesthesia in Pacific oyster. *Crassostrea gigas. Aquat. Living Resour.* 22, 29–34. doi: 10.1051/alr/2009006

Suquet, M., Araya, R. G., Lebrun, L., Queau, I., Mingant, C., and Robert, R. (2010). Anaesthesia and gonad sampling in the European flat oyster (*Ostrea edulis*). *Aquacult.* 308, 196–198. doi: 10.1016/j.aquaculture.2010.09.001

Taslima, K., Davie, A., Mcandrew, B. J., and Penman, D. J. (2016). DNA sampling from mucus in the Nile tilapia, *Oreochromis niloticus*: minimally invasive sampling for aquaculture-related genetics research. *Aquacult. Res.* 47, 4032–4037. doi: 10.1111/are.12809

Thomsen, P. F., and Willerslev, E. (2015). Environmental DNA–An emerging tool in conservation for monitoring past and present biodiversity. *Biol. Conserv.* 183, 4–18. doi: 10.1016/j.biocon.2014.11.019

Waldbusser, G. G., Hales, B., Langdon, C. J., Haley, B. A., Schrader, P., Brunner, E. L., et al. (2015). Saturation-state sensitivity of marine bivalve larvae to ocean acidification. *Nat. Climate Change* 5, 273–280. doi: 10.1038/nclimate2479

Walsh, P. S., Metzger, D. A., and Higuchi, R. (1991). Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques* 10 (4), 506–513.

Yanick, J. F., and Heath, D. D. (2000). Survival and growth of mussels subsequent to hemolymph sampling for DNA. *J. Shellfish Res.* 19, 991–993.

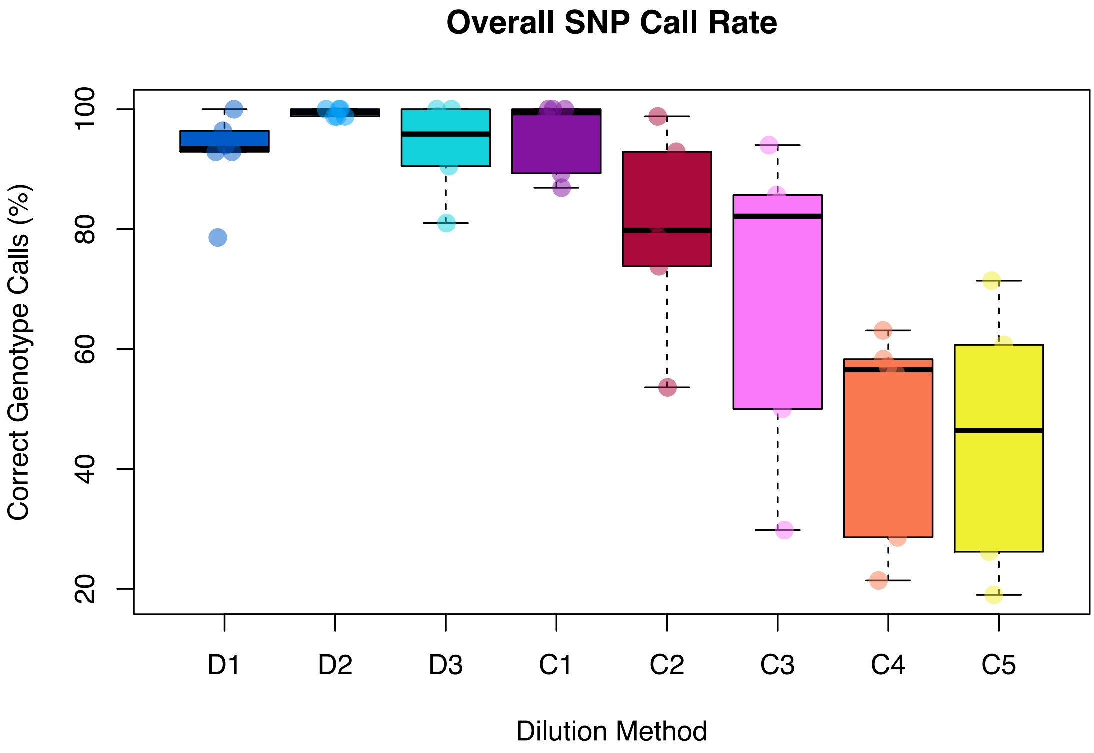
Conflict of Interest: LH is supported through a CASE studentship by Xelect Ltd. IJ, TA, and CH are employees of Xelect Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Holman, Hollenbeck, Ashton and Johnston. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

5.2 Supplementary Information

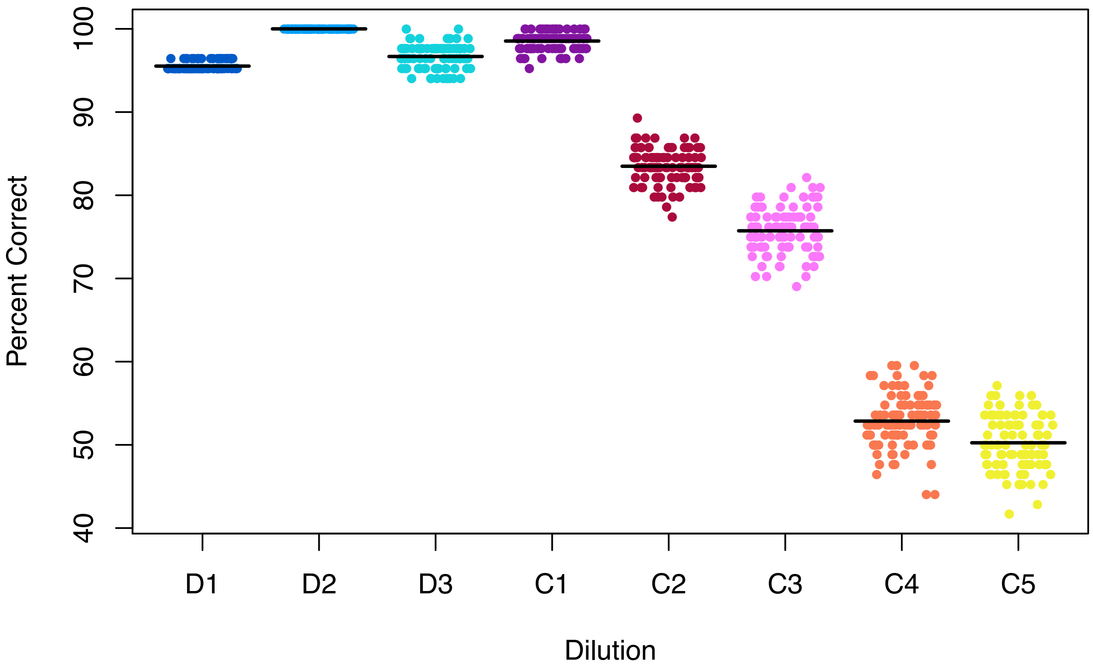
5.2.1 Supplementary Figure 1

Boxplot detailing overall proportion of eDNA derived SNP genotype calls compared to tissue derived calls for each Ostrea edulis individual. Each coloured box shows results for the dilutions as detailed in the main manuscript. The central black of each box is the median value, the box edges are the 25th and 75th percentiles and the outlier whiskers indicate the nearest value within 1.5 times the interquartile range. Raw data points have been overlaid on each box.



5.2.2 Supplementary Figure 2

Scatterplot detailing percentage of correct environmental DNA derived genotype calls in comparison to the tissue extractions for the eight dilutions detailed in the main manuscript. Each point represents the total result derived from a random sample of two replicates per individual Ostrea edulis genotype. The black line indicates the average percentage correct eDNA genotypes across the 100 random samples



Chapter 6 Synthesis & Conclusions

6.1 Overview

This thesis demonstrates some of the ways in which eDNA can contribute to our knowledge of marine biodiversity and biological invasions. As in Fig. 1 below, each chapter can be subdivided into fundamental and methodological findings relevant to the use of eDNA in marine biodiversity and invasion science.

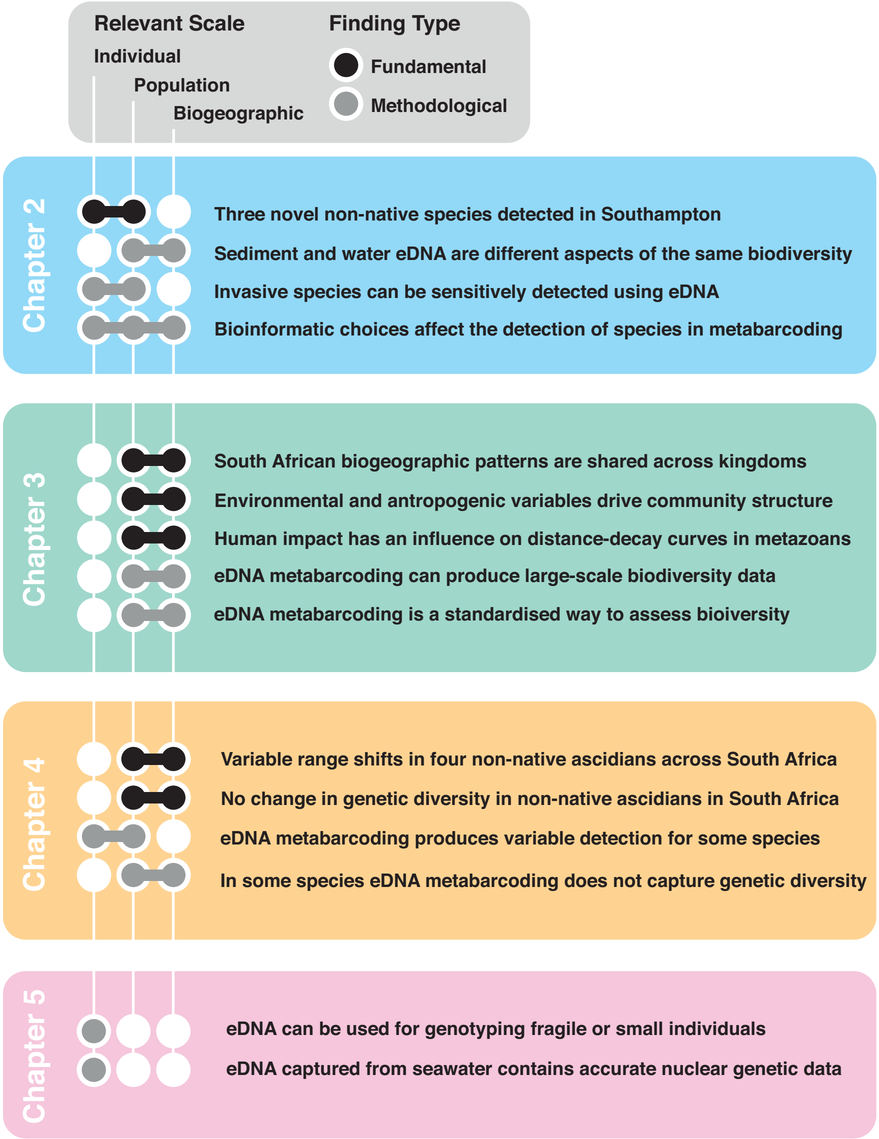

Relevant Scale
Individual
Population
Biogeographic
Finding Type
Fundamental
Methodological
Chapter 2
Three novel non-native species detected in Southampton
Sediment and water eDNA are different aspects of the same biodiversity
Invasive species can be sensitively detected using eDNA
Bioinformatic choices affect the detection of species in metabarcoding
Chapter 3
South African biogeographic patterns are shared across kingdoms
Environmental and antropogenic variables drive community structure
Human impact has an influence on distance-decay curves in metazoans
eDNA metabarcoding can produce large-scale biodiversity data
eDNA metabarcoding is a standardised way to assess bioiversity
Chapter 4
Variable range shifts in four non-native ascidians across South Africa
No change in genetic diversity in non-native ascidians in South Africa
eDNA metabarcoding produces variable detection for some species
In some species eDNA metabarcoding does not capture genetic diversity
Chapter 5
eDNA can be used for genotyping fragile or small individuals
eDNA captured from seawater contains accurate nuclear genetic data

Figure 1 Principal thesis findings and relevance at different spatial scales. Each filled circle shows the scale at which each finding is relevant, with findings in fundamental marine biodiversity or invasion biology shown in black and methodological or technical findings in grey.

In **Chapter Two** sediment and water samples collected from marinas across the UK were subject to eDNA metabarcoding of 18S and COI gene fragments. Three novel non-native species were detected: *Arcuatula senhousia*, *Cephalothrix simula* and *Paranais frici.* This demonstrates the sensitivity and utility of eDNA for the early detection on non-native species during an introduction. The eDNA from sediment and water were found to contain significantly different biodiversity within the same site, with differences between sites and between samples within a site having approximately equivalent community dissimilarity. Also, the detection of some species were found to depend on the parameters employed in bioinformatic pipelines, with reads corresponding with the non-native bryozoan *Bugula neritina* filtered out when applying quality thresholds typical of most eDNA metabarcoding studies. Overall, Chapter Two showed eDNA metabarcoding to be an reliable method for the detection of non-native species, and established appropriate protocols for further chapters.

In **Chapter Three** water samples were collected from 18 sites across the coastline of South Africa and subject to eDNA extraction and metabarcoding at 16S, 18S and COI gene fragments. The data were then subset into animal, protist and bacteria datasets for subsequent analysis. Across the different taxonomic groups consistent patterns of beta diversity were found along the South African coastline, with samples clustering with previously described ecoregions. The effect of anthropogenic and environmental parameters on patterns of beta diversity were evaluated; a minority of the observed variance between samples could be explained by temperature and anthropogenic impact for metazoans and bacteria, with chlorophyll *a* and anthropogenic impact having greater influence in the protist dataset. Finally, by comparing the community similarity to the distance between sites of varying human impact, evidence supported regional homogenisation of artificial sites in metazoans. This chapter provided an in-depth examination of the consistency of biodiversity patterns across taxa, demonstrating that eDNA can be used for standardised biodiversity assessments across large geographic distances and across the taxonomic tree of life.

Chapter Four explored the distribution and genetics of four non-native ascidians in South Africa. Rapid assessment data was collected and compared to circa. 2009 surveys. These changes in geographic range were then compared to regional historical distribution records. A mixed picture emerged, with stable range extent, range expansions and contractions found across the four species. A region of the COI gene was sequenced for a representative number of individuals for all species which showed variation across the coastline and within samples, but no significant difference between sampling occasions. Species incidence data from water eDNA metabarcoding showed variable accuracy compared to rapid assessment surveys. Data from the 18S region agreed with rapid assessment surveys, but did not detect two of the four species. In contrast COI data detected all four species, but showed stochastic detection, with detections where some rapid assessment results showed no detection and non-detections for species in some sites where rapid assessment surveys had a positive detection. The tissue derived COI sequences showed a large number of haplotypes not found in the COI eDNA metabarcoding data. However, the COI eDNA data did capture the most common haplotype in almost all species and sites. Overall, Chapter Four showed how non-native species shift their ranges over time and revealed limited turnover of haplotypes over almost a decade. The eDNA metabarcoding results showed some broad agreement with the observations of existing tools, but did not recover fine scale incidence and genetic data.

In **Chapter Five** a method for non-invasive genotyping using eDNA was presented and applied to a mollusk species of aquaculture interest. The eDNA samples showed very accurate nuclear SNP genotypes in comparison to tissue derived DNA samples. This demonstrated that eDNA contains nuclear genetic information for individuals, and that, for some genotyping platforms, correct genotyping relies on minimal dilution of eDNA. The study showed that eDNA genotyping of small or fragile individuals is both possible, and useful for aquaculture. The results are broadly relevant to any filter feeding species and question related to aquatic organisms.

6.2 Synthesis

6.2.1 A revolution in marine ecosystem monitoring

Anthropocene oceans are under threat and accurate, sensitive tools need to be developed and deployed globally to enable standardised marine monitoring (McGill et al. 2015, Dornelas et al. 2019, Bowler et al. 2020, Duarte et al. 2020). The recent, rapid rise of DNA-based biomonitoring tools shows huge potential for improving marine ecosystem management and research (Aylagas et al. 2020, Cordier et al. 2020). This thesis highlights some of the key strengths and weaknesses of this approach for marine biomonitoring and biodiversity research.

A growing body of literature demonstrates that eDNA-based methods have greater detection sensitivity than existing methods for generating biodiversity data (Deiner et al. 2017, Fediajevaite et al. 2021). Chapter Two and Four showed that eDNA metabarcoding detected some species missed by existing methods, though, in both chapters, there were also species found by existing methods that eDNA metabarcoding did not detect. It is accepted that some metabarcoding primers do not amplify certain taxonomic groups (Leray et al. 2013, Elbrecht et al. 2019) and primer development for accurate marine biomonitoring is ongoing (Wangensteen et al. 2018, Komai et al. 2019). The remaining false-negative eDNA detections are likely to be related to methodological limitations of current metabarcoding protocols and research continues into the appropriate depth of sequencing (Singer et al. 2019), choice of primers (Deiner et al. 2017, Collins et al. 2019) and the effects of laboratory choices when constructing and sequencing eDNA (Leray and Knowlton 2017, Nichols et al. 2018, Beentjes et al. 2019).

An advantage of eDNA metabarcoding is the ability to provide biodiversity observations for species that are yet to be described in a standardised way. Recent estimates find up to 37% of total species records in marine surveys may be undescribed by science (Appeltans et al. 2012) making biodiversity assessments difficult. Chapter Three used a large number of observations from sequences assigned to kingdom level to explore biogeographic patterns. These analyses demonstrate that ecological information can be derived from metabarcoding data even if species level assignments are not possible due to lack of DNA reference data (Weigand et al. 2019, Locatelli et al. 2020). Recent global scale analyses of genomic data collected from environmental samples have revealed vast undescribed marine biodiversity (de Vargas et al. 2015, Sunagawa et al. 2020). DNA-based tools that allow us to monitor and understand these as-of-yet undiscovered species are therefore essential for marine biodiversity researchers and managers.

There is increasing recognition of the importance of protecting intraspecific genetic diversity, as well as species biodiversity (Des Roches et al. 2018, Des Roches et al. 2021). However, collecting tissue samples of marine species, particularly in remote ocean environments, can be challenging and expensive (Nygård et al. 2016). Chapter Four demonstrated that eDNA metabarcoding data, collected with a view to capture species diversity, also captured broad-scale haplotype variation for some species. Chapter Five showed that eDNA contains nuclear genomic information and that accurate detection depends on methodological optimisation. Together, these observations support the growing body of literature suggesting that eDNA-based pan-organismal genetic diversity monitoring is possible in marine systems (Sigsgaard et al. 2020, Turon et al. 2020).

Despite the encouraging results of eDNA-based methods, there remains value in long-term ecological datasets to detect trends not apparent at smaller temporal scales (Edwards et al. 2010). We should be cautious of abandoning existing marine monitoring tools and use them to gather additional information not provided by eDNA-based methods (for example, age structure from fish trawling data). Yet, eDNA-based tools remain an increasingly important part of marine research and management, when deployed alongside existing methods for biodiversity monitoring.

6.2.2 Non-native species and eDNA

Further introductions of species are predicted well into the 21st Century (Seebens et al. 2021), therefore methods to assist in management and research of these species are vital. Some of the earliest eDNA studies targeted non-native species (Ficetola et al. 2008, Jerde et al. 2011) and the use of DNA-based methods for early detection and management have rapidly increased (Duarte et al. 2021).

Chapter Two found that eDNA metabarcoding was able to detect three previously unreported non-native species. These incidental detections are interesting, and of value to managers, but are not part of any formal monitoring network: researchers may choose different quality control thresholds, raising the possibility of a false-positive or detection with incorrect taxonomic assignment (Darling et al. 2020). The stochastic detections of non-native ascidians in Chapter Four also demonstrate that current eDNA metabarcoding protocols may fail to detect new arrivals, giving managers and researchers a false impression of introduction rates. Full eDNA metabarcoding-based surveillance facilitating early detection of marine non-native species will require standardised reporting, quality assurance and a DNA reference database with vouchered specimens. Implementing a large-scale biomonitoring scheme would therefore be expensive and technically challenging (Elbrecht and Steinke 2019, Piper et al. 2019). However, the increasing global economic impact of non-native species, and decreased cost of management when introductions are intercepted early, make surveillance programs a pragmatic option (Leung et al. 2002, Diagne et al. 2021).

Communication with, and between stakeholders, agencies and the broader public remains an issue in the use of eDNA monitoring for the management of non-native species. The detection of eDNA from a non-native species does not necessary indicate presence in the environment and if live organisms evade physical capture uncertainty can cause profound difficulties [such as in the case of Asian carp in the Great Lakes, North America(Jerde 2021)]. Conversely, as outlined above and in Chapters Two and Three, a negative detection is possible even when the target species is found in the environment. Recent work has suggested a framework (Mosher et al. 2020) and a decision tree (Sepulveda et al. 2020) to assist with communication and decision making regarding the use of DNA evidence in natural resource management.

6.2.3 Primer choice and eDNA metabarcoding

Research has consistently identified that different primers used for metabarcoding produce datasets with variable properties (Wangensteen et al. 2018, Collins et al. 2019, Elbrecht et al. 2019). The most apparent difference is the taxonomic composition and richness of biodiversity recovered, with broad target primer sets detecting much greater taxonomic diversity and richness compared to those that target a narrower taxonomic composition (Leray et al. 2013, Zhan et al. 2013, Miya et al. 2020). However, a lesser known distinction between these types of primer sets is the reliability between replicates and stochasticity of detection, with primers that amplify a smaller number of targets in a mixed DNA samples producing greater reliability and less stochastic detections (Collins et al. 2019).

Fig. 2 below shows three hypothetical sets of primers that are representative of the different primers used in eDNA metabarcoding experiments following Kelly et al. (2019). The first type (Fig. 2a) are broad target (sometimes called 'universal') primers that capture diversity across many different taxa, amplifying few taxa very well and few taxa very poorly. They are typically designed based on previously developed primers used for DNA barcoding (Leray et al. 2013, Wangensteen et al. 2018). The second type (Fig. 2b) amplify most taxa poorly and a few taxa well, for example Kelly et al. (2016). A third type (Fig. 2c) amplify a small number of taxa very well and most poorly if at all. These primers typically target a specific group (for example, fish or vertebrates) and a smaller number of species (Miya et al. 2020).

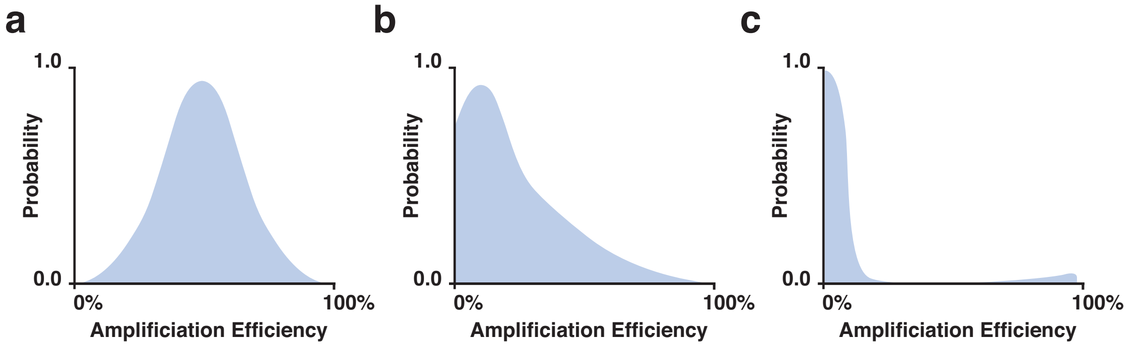


Figure 2 Distribution of amplification efficiencies for three hypothetical sets of primers, adapted from Kelly et al. (2019). **a** amplification efficiencies across a typical mixed DNA sample for a broad diversity primer set such as (Leray et al. 2013). **b** amplification efficiencies across a typical mixed DNA sample for a right-skewed primer set such as Kelly et al. (2016). **c** amplification efficiencies across a typical mixed DNA sample for a taxonomically specific primer set such as Miya et al. (2020)

The COI primers used in Chapters Two, Three and Four (Leray et al. 2013) fit into the first category (Fig. 2a). Many species were accurately detected in Chapter Two, and broad scale biogeographic patterns were identified in Chapter Three. However, the stochastic detection of non-native ascidians in Chapter Four showed that these primers are not suitable for all applications. Similar patterns were found with the 18S primers used in Chapter Two, Three and Four (Zhan et al. 2013), but with less taxonomic richness in comparison to the COI primers and less stochastic detections in Chapter Four. The amplification efficiency for these primers is likely between those shown in Fig. 2a & 2b.

All of these types of primers can produce useful ecological information, but some may be more suitable for specific questions than others. For example, a marine ecosystem manager may be interested in the broad-scale response of ecological communities to an environmental stressor. In this case, the stochastic detection found in some primer sets (Fig. 2a) would not obscure the whole community response, and the broad taxonomic composition profiled would be helpful to understand responses across different taxa. In contrast, a manager interested in the habitat use of a number of closely related shark species would find the reliability and accuracy found in primer sets with limited taxonomic scope (Fig. 2c) more useful.

Some authors have compared detection rates from broad target metabarcoding primers (such as Fig. 2a) to qPCR, claiming that metabarcoding offers less reliable results (Harper et al. 2018, Bylemans et al. 2019, Blackman et al. 2020). Since qPCR primers have an amplification profile similar to that of Fig. 2c, these comparisons reflect differences of primer choice, rather than an inherent difference of the detection method (high throughput sequencing versus quantitative PCR). Recent work has shown that eDNA metabarcoding with primers targeting a single species can provide greater sensitivity than qPCR in some cases (Westfall et al. 2021).

6.3 Future Work

This thesis explored how eDNA can contribute to knowledge of marine biodiversity and biological invasions. Progress has been made on several key fundamental and methodological questions through the analysis of observational data. However, eDNA can also be used in time-series analyses and manipulative studies to explore causative links between environmental stressors and biological communities, as shown in the framework presented in Fig. 3 below.

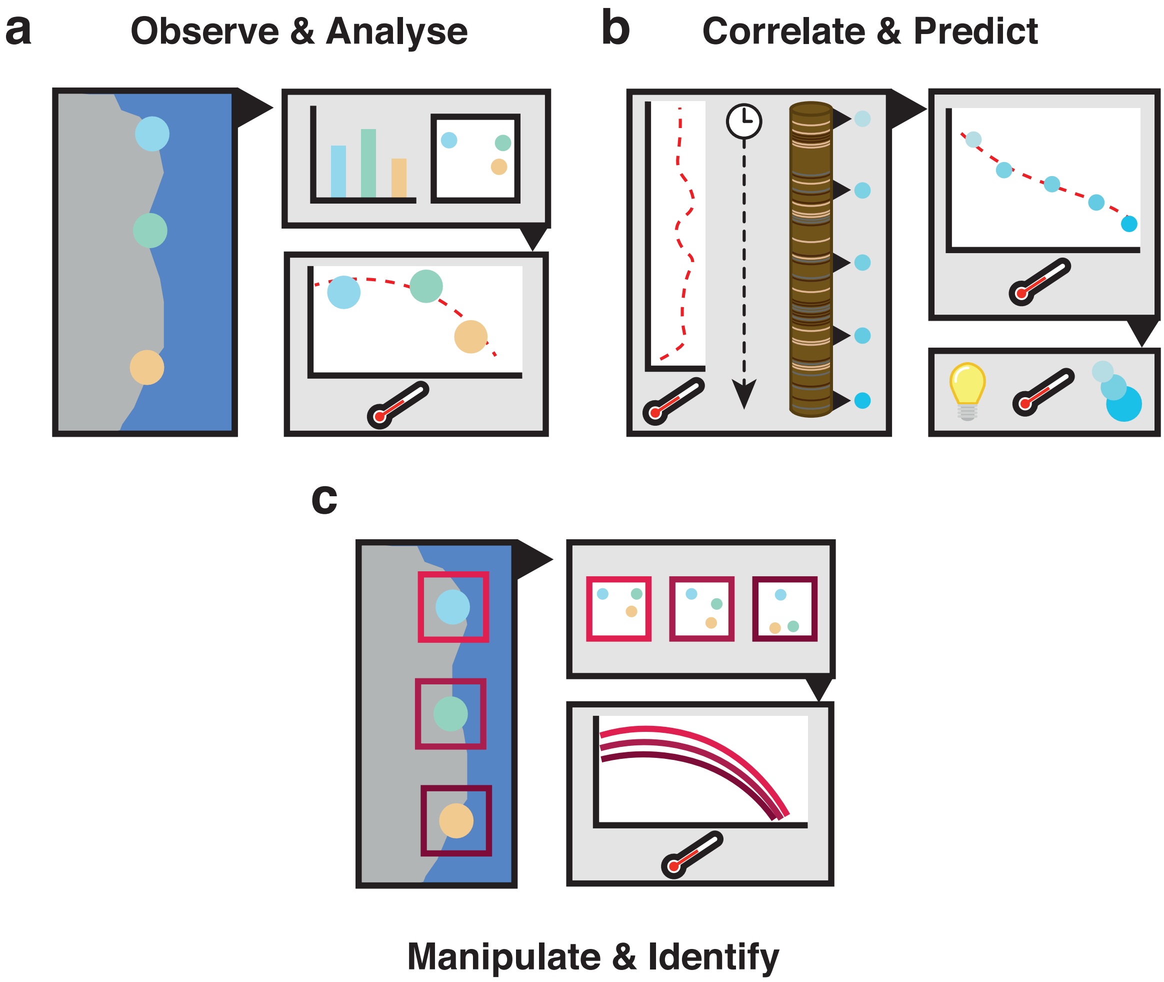


Figure 3 Conceptual framework showing how eDNA could be used in future studies to answer fundamental questions in invasion biology and marine biodiversity. **a** Observational studies using eDNA reveal high resolution ecological patterns that can be analysed to determine potential drivers (for example temperature). **b** Association studies using time series collected eDNA or ancient eDNA allow previous correlations to be corroborated and predictive models parameterised. **c** Manipulative studies in the laboratory use eDNA measurements to test models and identify a functional relationship between biological communities and the drivers affecting them.

In observation studies eDNA metabarcoding has been used to measure biodiversity and link these observations to simultaneously recorded environmental and biotic conditions (de Vargas et al. 2015, Sunagawa et al. 2015). A subsequent step would be to test these correlations using time-series data, or in other study systems, to produce a robust prediction of the underlying ecological mechanism. Time-series community data can be collected using eDNA samples stored in a biobank (Jarman et al. 2018) or through the use of sedimentary ancient eDNA (Pedersen et al. 2015, Pedersen et al. 2016). Environmental proxies and historical records can be used to resolve enviromental parameters of interest (Balint et al. 2018, Ellegaard et al. 2020). Finally the validated relationship from time-series data can be functionally identified using eDNA measured community metrics from manipulative experiments either in the lab or field (Nascimento et al. 2019). This framework takes advantage of the key methodological features of eDNA-based methods; standardisation across taxa, accuracy and sensitivity.

As an example, if we apply this framework to the Chapter Three we might first aim to test if the effect of sea surface temperature on community dissimilarity occurs persistently across time and space. This could be achieved through eDNA metabarcoding of sediment archives from the study area to test if temperature changes over time correlate with community changes. Alternativly, the study could be repeated across other temperature gradients to check for a consistent community response in other ecosystems. From these experiments we would hope to gain a functional prediction of the way in which sea temperature affects communities. Manipulative experiments could artificially heat or cool biological communities, either in the laboratory (Wernberg et al. 2012) or *in situ* (Ashton et al. 2017). The collection of biodiversity data from these experiments using eDNA-based techniques could then be used to evaluate causative links between temperature and community diversity metrics across taxa.

Not all biological questions can be easily decomposed into hypotheses suitable for investigation in this way. Additionally, some findings might not support continued ‘forward’ investigation in the framework but instead warrant further observational or correlational studies. These limitations aside eDNA-based methods clearly have utility for future investigation into the drivers and consequences of marine biodiversity change, and are broadly useful for biodiversity researchers and managers globally.

6.4 References

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

Ashton, G. V., S. A. Morley, D. K. A. Barnes, M. S. Clark, and L. S. Peck. 2017. Warming by 1°C Drives Species and Assemblage Level Responses in Antarctica's Marine Shallows. Current Biology **27**:2698-2705.e2693.

Aylagas, E., A. Borja, X. Pochon, A. Zaiko, N. Keeley, K. Bruce, P. Hong, G. M. Ruiz, E. D. Stein, S. Theroux, N. Geraldi, A. Ortega, L. Gajdzik, D. J. Coker, Y. Katan, T. Hikmawan, A. Saleem, S. Alamer, B. H. Jones, C. M. Duarte, J. Pearman, and S. Carvalho. 2020. Translational Molecular Ecology in practice: Linking DNA-based methods to actionable marine environmental management. Science of the Total Environment **744**:140780.

Balint, M., M. Pfenninger, H. P. Grossart, P. Taberlet, M. Vellend, M. A. Leibold, G. Englund, and D. Bowler. 2018. Environmental DNA time series in ecology. Trends in Ecology & Evolution **33**:945-957.

Beentjes, K. K., A. G. Speksnijder, M. Schilthuizen, M. Hoogeveen, and B. B. van der Hoorn. 2019. The effects of spatial and temporal replicate sampling on eDNA metabarcoding. PeerJ **7**:e7335.

Blackman, R. C., K. K. S. Ling, L. R. Harper, P. Shum, B. Hanfling, and L. Lawson-Handley. 2020. Targeted and passive environmental DNA approaches outperform established methods for detection of quagga mussels, *Dreissena rostriformis bugensis* in flowing water. Ecology and Evolution **10**:13248-13259.

Bowler, D. E., A. D. Bjorkman, M. Dornelas, I. H. Myers-Smith, L. M. Navarro, A. Niamir, S. R. Supp, C. Waldock, M. Winter, M. Vellend, S. A. Blowes, K. Böhning-Gaese, H. Bruelheide, R. Elahi, L. H. Antão, J. Hines, F. Isbell, H. P. Jones, A. E. Magurran, J. S. Cabral, and A. E. Bates. 2020. Mapping human pressures on biodiversity across the planet uncovers anthropogenic threat complexes. People and Nature **2**:380-394.

Bylemans, J., D. M. Gleeson, R. P. Duncan, C. M. Hardy, and E. M. Furlan. 2019. A performance evaluation of targeted eDNA and eDNA metabarcoding analyses for freshwater fishes. Environmental DNA **1**:402-414.

Collins, R. A., J. Bakker, O. S. Wangensteen, A. Z. Soto, L. Corrigan, D. W. Sims, M. J. Genner, and S. Mariani. 2019. Non-specific amplification compromises environmental DNA metabarcoding with COI. Methods in Ecology and Evolution **10**:1985-2001.

Cordier, T., L. Alonso-Sáez, L. Apothéloz-Perret-Gentil, E. Aylagas, D. A. Bohan, A. Bouchez, A. Chariton, S. Creer, L. Frühe, F. Keck, N. Keeley, O. Laroche, F. Leese, X. Pochon, T. Stoeck, J. Pawlowski, and A. Lanzén. 2020. Ecosystems monitoring powered by environmental genomics: A review of current strategies with an implementation roadmap. Molecular Ecology **In Press**.

Darling, J. A., X. Pochon, C. L. Abbott, G. J. Inglis, and A. Zaiko. 2020. The risks of using molecular biodiversity data for incidental detection of species of concern. Diversity and Distributions **26**:1116-1121.

de Vargas, C., S. Audic, N. Henry, J. Decelle, F. Mahé, R. Logares, E. Lara, C. Berney, N. Le Bescot, I. Probert, M. Carmichael, J. Poulain, S. Romac, S. Colin, J.-M. Aury, L. Bittner, S. Chaffron, M. Dunthorn, S. Engelen, O. Flegontova, L. Guidi, A. Horák, O. Jaillon, G. Lima-Mendez, J. Lukeš, S. Malviya, R. Morard, M. Mulot, E. Scalco, R. Siano, F. Vincent, A. Zingone, C. Dimier, M. Picheral, S. Searson, S. Kandels-Lewis, S. G. Acinas, P. Bork, C. Bowler, G. Gorsky, N. Grimsley, P. Hingamp, D. Iudicone, F. Not, H. Ogata, S. Pesant, J. Raes, M. E. Sieracki, S. Speich, L. Stemmann, S. Sunagawa, J. Weissenbach, P. Wincker, and E. Karsenti. 2015. Eukaryotic plankton diversity in the sunlit ocean. Science **348**:1261605.

Deiner, K., H. M. Bik, E. Machler, M. Seymour, A. Lacoursiere-Roussel, F. Altermatt, S. Creer, I. Bista, D. M. Lodge, N. de Vere, M. E. Pfrender, and L. Bernatchez. 2017. Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. Molecular Ecology **26**:5872-5895.

Des Roches, S., L. H. Pendleton, B. Shapiro, and E. P. Palkovacs. 2021. Conserving intraspecific variation for nature's contributions to people. Nature Ecology & Evolution.

Des Roches, S., D. M. Post, N. E. Turley, J. K. Bailey, A. P. Hendry, M. T. Kinnison, J. A. Schweitzer, and E. P. Palkovacs. 2018. The ecological importance of intraspecific variation. Nature Ecology & Evolution **2**:57-64.

Diagne, C., B. Leroy, A.-C. Vaissière, R. E. Gozlan, D. Roiz, I. Jarić, J.-M. Salles, C. J. A. Bradshaw, and F. Courchamp. 2021. High and rising economic costs of biological invasions worldwide. Nature **592**:571-576.

Dornelas, M., E. M. P. Madin, M. Bunce, J. D. DiBattista, M. Johnson, J. S. Madin, A. E. Magurran, B. J. McGill, N. Pettorelli, O. Pizarro, S. B. Williams, M. Winter, and A. E. Bates. 2019. Towards a macroscope: Leveraging technology to transform the breadth, scale and resolution of macroecological data. Global Ecology and Biogeography **28**:1937-1948.

Duarte, C. M., S. Agusti, E. Barbier, G. L. Britten, J. C. Castilla, J.-P. Gattuso, R. W. Fulweiler, T. P. Hughes, N. Knowlton, C. E. Lovelock, H. K. Lotze, M. Predragovic, E. Poloczanska, C. Roberts, and B. Worm. 2020. Rebuilding marine life. Nature **580**:39-51.

Duarte, S., P. E. Vieira, A. S. Lavrador, and F. O. Costa. 2021. Status and prospects of marine NIS detection and monitoring through (e)DNA metabarcoding. Science of the Total Environment **751**:141729.

Edwards, M., G. Beaugrand, G. C. Hays, J. A. Koslow, and A. J. Richardson. 2010. Multi-decadal oceanic ecological datasets and their application in marine policy and management. Trends in Ecology & Evolution **25**:602-610.

Elbrecht, V., T. W. Braukmann, N. V. Ivanova, S. W. Prosser, M. Hajibabaei, M. Wright, E. V. Zakharov, P. D. Hebert, and D. Steinke. 2019. Validation of COI metabarcoding primers for terrestrial arthropods. PeerJ **7**:e7745.

Elbrecht, V., and D. Steinke. 2019. Scaling up DNA metabarcoding for freshwater macrozoobenthos monitoring. Freshwater Biology **64**:380-387.

Ellegaard, M., R. J. Clokie Martha, T. Czypionka, D. Frisch, A. Godhe, A. Kremp, A. Letarov, T. J. McGenity, S. Ribeiro, and N. John Anderson. 2020. Dead or alive: sediment DNA archives as tools for tracking aquatic evolution and adaptation. Communications Biology **3**:1.

Fediajevaite, J., V. Priestley, R. Arnold, and V. Savolainen. 2021. Meta-analysis shows that environmental DNA outperforms traditional surveys, but warrants better reporting standards. Ecology and Evolution **11**:4803-4815.

Ficetola, G. F., C. Miaud, F. Pompanon, and P. Taberlet. 2008. Species detection using environmental DNA from water samples. Biology Letters **4**:423-425.

Harper, L. R., L. Lawson Handley, C. Hahn, N. Boonham, H. C. Rees, K. C. Gough, E. Lewis, I. P. Adams, P. Brotherton, S. Phillips, and B. Hänfling. 2018. Needle in a haystack? A comparison of eDNA metabarcoding and targeted qPCR for detection of the great crested newt (*Triturus cristatus*). Ecology and Evolution **8**:6330-6341.

Jarman, S. N., O. Berry, and M. Bunce. 2018. The value of environmental DNA biobanking for long-term biomonitoring. Nature Ecology and Evolution **2**:1192-1193.

Jerde, C. L. 2021. Can we manage fisheries with the inherent uncertainty from eDNA? Journal of Fish Biology **98**:341-353.

Jerde, C. L., A. R. Mahon, W. L. Chadderton, and D. M. Lodge. 2011. "Sight-unseen" detection of rare aquatic species using environmental DNA. Conservation Letters **4**:150-157.

Kelly, R. P., J. L. O'Donnell, N. C. Lowell, A. O. Shelton, J. F. Samhouri, S. M. Hennessey, B. E. Feist, and G. D. Williams. 2016. Genetic signatures of ecological diversity along an urbanization gradient. PeerJ **4**:e2444.

Kelly, R. P., A. O. Shelton, and R. Gallego. 2019. Understanding PCR Processes to Draw Meaningful Conclusions from Environmental DNA Studies. Scientific Reports **9**:12133.

Komai, T., R. O. Gotoh, T. Sado, and M. Miya. 2019. Development of a new set of PCR primers for eDNA metabarcoding decapod crustaceans. Metabarcoding and Metagenomics **3**:e33835.

Leray, M., and N. Knowlton. 2017. Random sampling causes the low reproducibility of rare eukaryotic OTUs in Illumina COI metabarcoding. PeerJ **5**:e3006.

Leray, M., J. Y. Yang, C. P. Meyer, S. C. Mills, N. Agudelo, V. Ranwez, J. T. Boehm, and R. J. Machida. 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.

Leung, B., D. M. Lodge, D. Finnoff, J. F. Shogren, M. A. Lewis, and G. Lamberti. 2002. An ounce of prevention or a pound of cure: bioeconomic risk analysis of invasive species. Proceedings of the Royal Society of London. Series B: Biological Sciences **269**:2407-2413.

Locatelli, N. S., P. B. McIntyre, N. O. Therkildsen, and D. S. Baetscher. 2020. GenBank's reliability is uncertain for biodiversity researchers seeking species-level assignment for eDNA. Proceedings of the National Academy of Sciences **117**:32211.

McGill, B. J., M. Dornelas, N. J. Gotelli, and A. E. Magurran. 2015. Fifteen forms of biodiversity trend in the Anthropocene. Trends in Ecology & Evolution **30**:104-113.

Miya, M., R. O. Gotoh, and T. Sado. 2020. MiFish metabarcoding: a high-throughput approach for simultaneous detection of multiple fish species from environmental DNA and other samples. Fisheries Science **86**:939–970.

Mosher, B. A., R. F. Bernard, J. M. Lorch, D. A. W. Miller, K. L. D. Richgels, C. L. White, and E. H. Campbell Grant. 2020. Successful molecular detection studies require clear communication among diverse research partners. Frontiers in Ecology and the Environment **18**:43-51.

Nascimento, F. J. A., M. Dahl, D. Deyanova, L. D. Lyimo, H. M. Bik, T. Schuelke, T. J. Pereira, M. Björk, S. Creer, and M. Gullström. 2019. Above-below surface interactions mediate effects of seagrass disturbance on meiobenthic diversity, nematode and polychaete trophic structure. Communications Biology **2**:362.

Nichols, R. V., C. Vollmers, L. A. Newsom, Y. Wang, P. D. Heintzman, M. Leighton, R. E. Green, and B. Shapiro. 2018. Minimizing polymerase biases in metabarcoding. Molecular Ecology Resources **18**:927-939.

Nygård, H., S. Oinonen, H. A. Hällfors, M. Lehtiniemi, E. Rantajärvi, and L. Uusitalo. 2016. Price vs. Value of Marine Monitoring. Frontiers in Marine Science **3**:205.

Pedersen, M. W., S. Overballe-Petersen, L. Ermini, C. Der Sarkissian, J. Haile, M. Hellstrom, J. Spens, P. F. Thomsen, K. Bohmann, and E. Cappellini. 2015. Ancient and modern environmental DNA. Philosophical Transactions of the Royal Society B: Biological Sciences **370**:20130383.

Pedersen, M. W., A. Ruter, C. Schweger, H. Friebe, R. A. Staff, K. K. Kjeldsen, M. L. Mendoza, A. B. Beaudoin, C. Zutter, and N. K. Larsen. 2016. Postglacial viability and colonization in North America's ice-free corridor. Nature **537**:45.

Piper, A. M., J. Batovska, N. O. I. Cogan, J. Weiss, J. P. Cunningham, B. C. Rodoni, and M. J. Blacket. 2019. Prospects and challenges of implementing DNA metabarcoding for high-throughput insect surveillance. GigaScience **8**.

Seebens, H., S. Bacher, T. M. Blackburn, C. Capinha, W. Dawson, S. Dullinger, P. Genovesi, P. E. Hulme, M. van Kleunen, I. Kühn, J. M. Jeschke, B. Lenzner, A. M. Liebhold, Z. Pattison, J. Pergl, P. Pyšek, M. Winter, and F. Essl. 2021. Projecting the continental accumulation of alien species through to 2050. Global Change Biology **27**:970-982.

Sepulveda, A. J., N. M. Nelson, C. L. Jerde, and G. Luikart. 2020. Are Environmental DNA Methods Ready for Aquatic Invasive Species Management? Trends in Ecology & Evolution **35**:668-678.

Sigsgaard, E. E., M. R. Jensen, I. E. Winkelmann, P. R. Møller, M. M. Hansen, and P. F. Thomsen. 2020. Population-level inferences from environmental DNA—Current status and future perspectives. Evolutionary Applications **13**:245-262.

Singer, G. A. C., N. A. Fahner, J. G. Barnes, A. McCarthy, and M. Hajibabaei. 2019. Comprehensive biodiversity analysis via ultra-deep patterned flow cell technology: a case study of eDNA metabarcoding seawater. Sci Rep **9**:5991.

Sunagawa, S., S. G. Acinas, P. Bork, C. Bowler, S. G. Acinas, M. Babin, P. Bork, E. Boss, C. Bowler, G. Cochrane, C. de Vargas, M. Follows, G. Gorsky, N. Grimsley, L. Guidi, P. Hingamp, D. Iudicone, O. Jaillon, S. Kandels, L. Karp-Boss, E. Karsenti, M. Lescot, F. Not, H. Ogata, S. Pesant, N. Poulton, J. Raes, C. Sardet, M. Sieracki, S. Speich, L. Stemmann, M. B. Sullivan, S. Sunagawa, P. Wincker, D. Eveillard, G. Gorsky, L. Guidi, D. Iudicone, E. Karsenti, F. Lombard, H. Ogata, S. Pesant, M. B. Sullivan, P. Wincker, C. de Vargas, and C. Tara Oceans. 2020. Tara Oceans: towards global ocean ecosystems biology. Nature Reviews Microbiology **18**:428-445.

Sunagawa, S., L. P. Coelho, S. Chaffron, J. R. Kultima, K. Labadie, G. Salazar, B. Djahanschiri, G. Zeller, D. R. Mende, A. Alberti, F. M. Cornejo-Castillo, P. I. Costea, C. Cruaud, F. d'Ovidio, S. Engelen, I. Ferrera, J. M. Gasol, L. Guidi, F. Hildebrand, F. Kokoszka, C. Lepoivre, G. Lima-Mendez, J. Poulain, B. T. Poulos, M. Royo-Llonch, H. Sarmento, S. Vieira-Silva, C. Dimier, M. Picheral, S. Searson, S. Kandels-Lewis, C. Bowler, C. de Vargas, G. Gorsky, N. Grimsley, P. Hingamp, D. Iudicone, O. Jaillon, F. Not, H. Ogata, S. Pesant, S. Speich, L. Stemmann, M. B. Sullivan, J. Weissenbach, P. Wincker, E. Karsenti, J. Raes, S. G. Acinas, and P. Bork. 2015. Structure and function of the global ocean microbiome. Science **348**:1261359.

Turon, X., A. Antich, C. Palacín, K. Præbel, and O. S. Wangensteen. 2020. From metabarcoding to metaphylogeography: separating the wheat from the chaff. Ecological Applications **30**:e02036.

Wangensteen, O. S., C. Palacin, M. Guardiola, and X. Turon. 2018. DNA metabarcoding of littoral hard-bottom communities: high diversity and database gaps revealed by two molecular markers. PeerJ **6**:e4705.

Weigand, H., A. J. Beermann, F. Čiampor, F. O. Costa, Z. Csabai, S. Duarte, M. F. Geiger, M. Grabowski, F. Rimet, B. Rulik, M. Strand, N. Szucsich, A. M. Weigand, E. Willassen, S. A. Wyler, A. Bouchez, A. Borja, Z. Čiamporová-Zaťovičová, S. Ferreira, K.-D. B. Dijkstra, U. Eisendle, J. Freyhof, P. Gadawski, W. Graf, A. Haegerbaeumer, B. B. van der Hoorn, B. Japoshvili, L. Keresztes, E. Keskin, F. Leese, J. N. Macher, T. Mamos, G. Paz, V. Pešić, D. M. Pfannkuchen, M. A. Pfannkuchen, B. W. Price, B. Rinkevich, M. A. L. Teixeira, G. Várbíró, and T. Ekrem. 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.

Wernberg, T., D. A. Smale, and M. S. Thomsen. 2012. A decade of climate change experiments on marine organisms: procedures, patterns and problems. Global Change Biology **18**:1491-1498.

Westfall, K. M., T. W. Therriault, and C. L. Abbott. 2021. Targeted Next Generation Sequencing of environmental DNA improves detection and quantification of invasive European green crab (*Carcinus maenas*). bioRxiv:2021.2003.2001.433397.

Zhan, A., M. Hulák, F. Sylvester, X. Huang, A. A. Adebayo, C. L. Abbott, S. J. Adamowicz, D. D. Heath, M. E. Cristescu, and H. J. MacIsaac. 2013. High sensitivity of 454 pyrosequencing for detection of rare species in aquatic communities. Methods in Ecology and Evolution **4**:558-565.