KU LEUVEN

FACULTY OF SCIENCE

Intercepting invaders: metabarcoding for monitoring non-indigenous species in a North Sea harbour

Anton BILSEN

Supervisor: Prof. F. Volckaert
Laboratory of Biodiversity and Evolutionary Genomics, KU Leuven

Co-supervisor: P. Hablützel
Flanders Marine Institute

Thesis presented in fulfillment of the requirements for the degree of Master of Science in 2021

Academic year 2020-2021

Acknowledgements

I wish to express my heartfelt thanks to all the staff at LBEG (Laboratory of Biodiversity and Evolutionary Genomics, KU Leuven) for listening to my presentations, making suggestions for my scripts, and helping with laboratory and bioinformatics work. Also, special thanks goes to Sander Delacauw from the Flanders Marine Institute for assisting me during the sampling campaign. My mentor, Pascal Hablützel, also deserves recognition for tirelessly helping me to plan and carry out the practical work as well as to proofread and comment on my manuscripts. Lastly, I would like to thank my family, without whose support this thesis would not have been possible.

Contents

Summary

Non-indigenous species (NIS) are a growing concern in coastal areas and have the potential to become invasive, displacing native organisms and harming economic activities. One major means whereby marine NIS are dispersed is via ballast water. Consequently, vessels are required by the International Maritime Organisation to treat their ballast water before entering a port. Exemptions to this requirement can, however, be granted if it can be shown that the voyage undertaken by the ship does not carry a significant risk of introducing new or dangerous alien species from one port to another. To evaluate the risk, a survey of the biota of all the harbours encountered during the voyage must be conducted. Biological surveys for NIS have traditionally relied on morphological identification of biota, but this approach is time- and labour-intensive and prone to errors or biases induced by a lack of taxonomic expertise. DNA-based survey methods, such as metabarcoding, have the potential to ameliorate these issues but are still not commonly implemented by stakeholders for the North Sea region.

The EU Interreg project GEANS (Global Ecosystem Health Assessment for the North Sea) aims to facilitate the implementation of genetic-molecular tools for routine ecosystem assessment in the North Sea region. The project consists of various pilot studies that aim to compare DNA-based and traditional survey methods of the marine environment. One of the research areas of GEANS is monitoring of NIS in harbours. My thesis formed part of a pilot study evaluating the performance of metabarcoding for detection of NIS in the harbour of Ostend (Belgium).

First, I compared the total numbers of native and non-native species detected from settlement plates via visual identification with the numbers reported from these plates via metabarcoding of the 18S rRNA gene. Next, I compared the performance of two marker genes (COI and 18S rRNA) in detecting species composition from plankton samples via metabarcoding. Metabarcoding was able to detect many more species from the settlement plates than morphological examination alone. It also enabled the rapid identification of planktonic species, which would otherwise be extremely difficult to identify morphologically. Non-native species were found from both planktonic and plate samples via metabarcoding. Some of these taxa were never recorded before in the North Sea while others were well-known invasive species. However, the lack of consensus seen in this study among the different species detection methods, coupled with many dubious identifications, indicate shortcomings of DNA-based methods. I demonstrated that metabarcoding may be feasible to detect NIS in North Sea harbours, but that current technical limitations of the method warrant caution and further explorative research.

Introduction

I. The problem of species introductions

The expansion of species ranges into areas previously unoccupied has always played an important role in shaping biogeographical processes throughout the history of the earth, as has been widely acknowledged. However, in recent years, increasing globalization and the concomitant drastic increase in both the distances covered by and the intensity of worldwide shipping and transport have artificially introduced many species into environments which they would never have been able to reach via natural migration or propagule spread (Kettunen et al., 2009; Seebens et al., 2013; Corrales et al., 2020). As a result, the number of species outside their natural range has increased by several orders of magnitude in the previous few decades (Seebens et al., 2013), especially in economically developed countries (Turbelin et al., 2017). Sometimes, the species in question (referred to as non-indigenous species, or NIS (Boudouresque & Verlaque, 2012; Cardeccia et al., 2018)) are a cause for real concern, both for the environments they are introduced into and for the human populations depending on those environments.

Most species introduced by humans outside their former, naturally colonized ranges are not able to establish free-living populations and, even if they do, often have no noticeable negative effects on the recipient ecosystem or on human populations (García-Berthou et al., 2005; Kettunen et al., 2009; Faulkner et al., 2014). In addition, many introduced organisms can be viewed as beneficial additions to their new ranges, both in terms of ecosystem functions and socio-economic interests. For example, NIS may provide food and shelter to endangered natives or offer a functional replacement of declining or extinct native taxa (Schlaepfer et al., 2011). Foreign, imported organisms also meet market demands for foodstuffs and ornamentals worldwide (Kettunen et al., 2009). Many other NIS, however, are able to thrive in their new ranges and do so at the expense of the native biota, displacing it through competition, predation, hybridization, or the introduction of new pathogens against which the native species have no resistance (Kettunen et al., 2009; Occhipinti-Ambrogi & Galil, 2010). The disruption of ecosystems that these foreign organisms bring about often proceeds to directly impact human economies and social well-being; examples include transmitting human diseases, causing allergies, destroying the appearance of culturally important sites, and compromising the health of economically important species (Kettunen et al., 2009). Especially prolific, harmful NIS are termed invasive alien species (IAS) (Turbelin et al., 2017; Cardeccia et al., 2018).

IAS have already been recognized as a cause of native species extinctions and major changes in ecosystem structure and functioning worldwide (Kettunen et al., 2009). IAS are, therefore, a major threat to global biodiversity, along with other anthropogenic stressors such as climate change, eutrophication and pollution (Occhipinti-Ambrogi, 2007; Kettunen et al., 2009; Occhipinti-Ambrogj & Galil, 2010; Seebens et al., 2013). IAS appearances are known to be facilitated by other aspects of global change such as climate warming and organic pollution, either by being directly favoured by the altered living conditions (such as warmer temperatures or higher nutrient concentrations, respectively) or by taking in ecological niches opened up when native species are no longer able to cope effectively with these stressors (Occhipinti-Ambrogi, 2007; Cheung et al., 2009; Karatayev et al., 2009; Occhipinti-Ambrogi & Galil, 2010; Corrales et al., 2020). Introduction rates of IAS show no signs of decline in the foreseeable future; on the contrary, they will probably continue to rise as global commerce intensifies (Sardain et al., 2019). IAS are notoriously difficult to eradicate once established, so that preventive measures are the only feasible option for avoiding these invasions (Faulkner et al., 2014; Ojaveer et al., 2018). However, preventive measures have historically shown varying levels of implementation and, therefore, mixed success (Sardain et al., 2019). Taking into account only two aspects of human-induced stress on the earth's system—climate change and invasive species—and the interactions between the two, Cheung et al. (2009) predicted a 60 % replacement of the current marine biodiversity by 2050.

Nonetheless, IAS have never commanded the same level of popular or academic interest as many of these other components of global change (Strayer et al., 2012; Ojaveer et al., 2018). This situation has in some cases led to a kind of vicious cycle where disregard by public and policymakers results in a lack of funding for IAS-related research; as a result, little information on the harmful impacts of IAS is available, leading to further disregard by public and policymakers (Ojaveer et al., 2018). Strayer (2012) notes that the study of biological invasions tends to be hampered by several biases and misconceptions, such as the assumption of "rules" with little observational evidence and the tendency to rely on so-called expert opinion based on arbitrary criteria of impact. While some taxonomic groups (such as invasive fish and molluscs) have received a considerable deal of attention within the discipline of invasion biology, other taxa (including many marine species such as barnacles, ascidians, bryozoans, and microbes) remain underrepresented in the literature on IAS (Corrales et al., 2020). In addition, predictive models tend to lack uncertainty estimates and information on validation procedures (Corrales et al., 2020), conditions which limit the application of these models. It is therefore imperative that invasion biologists adopt more rigorous, standardized methodology for not only describing current invasions but also predicting future ones, two crucial aspects of the discipline (Strayer, 2012; Ojaveer et al., 2018).

II. Defining NIS and IAS: different perspectives

II.1: NIS vs. IAS

The basic definitions of NIS and IAS have been briefly elucidated above, but their exact meanings, as well as those of related terms including *native*, *non-indigenous*, *invasive*, and *exotic*, tend to be used ambiguously (Gilroy et al., 2017). Different authors employ different definitions of the same terms, even though these terms represent fundamental concepts in invasion biology (Gilroy et al., 2017). The ambiguity in the academic literature has been reflected in policy regarding NIS, where countries set up their own criteria to use when determining if a species is non-native (and hence a potential threat) or native (Gilroy et al., 2017). This situation is unfortunate because the negative impacts of invasive NIS transcend political boundaries and hence require international cooperation to be controlled (Gilroy et al., 2017; Vitule et al., 2019).

Alien species (considered synonymous with introduced or exotic species by Webber & Scott, 2012) are generally taken to be species that occur outside their native ranges due to human-mediated dispersal (Corrales et al., 2020), regardless of whether these species exist under human care (such as animals in captivity) or have established free-living populations in their new ranges (Essl et al., 2018). NIS, in principle, are defined likewise (Boudoresque & Verlaque, 2012; Olenin et al., 2017; Corrales et al., 2020) and can thus be considered a synonym of alien species. Those NIS that exist as established populations, propagating without any human assistance, are termed introduced or naturalized species (Boudouresque & Verlaque, 2012; Essl et al., 2018).

NIS are often considered synonymously with IAS, as noted by Shrader-Frechette (2001), even though IAS strictly speaking are those NIS that, in addition to having established free-living populations, are especially prolific and/or exert a negative social or biological impact in their new ranges (Occhipinti-Ambrogi & Galil, 2010; Boudouresque & Verlaque, 2012; Turbelin et al., 2017; Cardeccia et al., 2018; Corrales et al., 2020). As noted in the introduction, many NIS do not have any noticeably harmful effects on their recipient ecosystems. It should also be noted that some native species may, as a result of artificial disturbance to their environment (such as the removal of predators or competitors), begin to rapidly proliferate and exert similar negative impacts to those of invasive aliens (Valéry et al., 2008). These organisms, however, are not considered IAS by Olenin et al. (2017), even though the difference with “true” IAS is more of a theoretical distinction than a biological or ecological one. Valéry et al. (2008) find this view logically inconsistent. Instead of focusing on species’ geographical origins, these authors focus on the concept of a biotic invasion itself, which they define based on events

occurring during the invasion process. In their view, all biotic invasions, whether of native or introduced organisms, result when a species' natural boundaries to proliferation are removed, allowing it to spread rapidly and giving it a competitive advantage relative to other species. This process-based definition further avoids the difficulty of having to decide what constitutes a sufficiently large, negative impact (Valéry et al., 2008) on the biota and/or human populations of the recipient ecosystems. Negative impact is indeed a somewhat subjective concept, according to Sagoff (2018), who is of the opinion that the very notion of an invasive species requires the use of ethical and moral arguments that are generally not within the reach of science. However, the multiple billions of euros spent on mitigating the effect of IAS worldwide and the general recognition of IAS being second only to habitat loss in threatening global biodiversity (Kettunen et al., 2009; Tollington et al., 2017) would suggest that these negative impacts are, at least for IAS as a whole, fairly obvious.

II.2: Determination of native range

The determination of an organism's native range is, however, not always straightforward. Geo-political boundaries are often considered when evaluating whether a species is indigenous to a given country or similar management unit, but this practice has led to considerable misrepresentations of NIS occurrence worldwide, especially in countries that span multiple ecoregions and contain multiple ecosystems (Essl et al., 2018; Vitule et al., 2019). Corrales et al. (2020) suggest that native range can be determined via biogeographic boundaries. These boundaries may be any physical or chemical environments that cannot normally be traversed alive by an organism and its propagules and thus prohibit range expansion beyond that point; examples include changes in water salinity and temperature in aquatic environments and mountain ranges and rivers in terrestrial habitats (Essl et al., 2018). Gilroy et al. (2017) expound upon the topic of dispersal barriers with the concept of the dispersal envelope, which gives the range that a species is expected to inhabit if it spreads using purely natural means. Appearances of the organism outside this range would therefore suggest human-mediated transport. Webber & Scott (2012) further elaborated this concept to define so-called projected dispersal envelopes, which give the expected distribution of an organism considering predictions regarding climate change and other anthropogenic stressors. However, disjunct distribution ranges and difficulties in accurately estimating dispersal capacity hamper the delineation of natural dispersal ranges for many species (Gilroy et al., 2017; Essl et al., 2018). The use of natural dispersal distance as a guide for assessing native range also requires that the forms of dispersal allowed within the definition be chosen beforehand, which Gilroy et al. (2017) consider to be an example of circular logic.

Another criterion that is routinely used in both the academic and the legislative literature to delineate a species' native range is the extent of its historical distribution, which is generally

taken to encompass the range inhabited by the organism before the advent of global trade networks (Gilroy et al., 2017). The start of globalization is often taken as a cut-off value for historical range because the majority of NIS have been spread as a result of the recent international trade, which has increased so rapidly that native systems have little time to adjust to the newcomers (Gilroy et al., 2017; Essl et al., 2018; Corrales et al., 2020). The impact of "modern" NIS, as opposed to species introduced by humans before globalization, can therefore be said to lie (at least in part) in the intensity with which they are introduced and the abundances in which they occur. However, some species are known to have been introduced before globalization and yet exert a negative influence on their new environments. Examples given by Ojaveer et al. (2018) include the mangrove-destroying isopod *Sphaeroma terebrans*, introduced into Brazil from the East Indies during the 19th century, and the cord grass *Spartina alterniflora*, a North American native found in France in 1803. Consequently, one cannot assume that an organism's historical range always equals its natural range.

II.3: Extent of human interference

Along with natural range, human-mediated dispersal is also a topic open to various interpretations. In many cases, foreign organisms are themselves traded and transported due to their economic value, and then either deliberately or accidentally released into the recipient ecosystem (Kettunen et al., 2009). At other times, species enter new environments as contaminants of goods or "hitch-hikers" on transportation vectors (such as ships or trucks) (Gilroy et al., 2017; Essl et al., 2018). Although trade in organisms (where a number of propagules escape or are released into the recipient area) certainly is an example of human-mediated dispersal, would the opening of a canal between two formerly disconnected regions (e.g., the Suez Canal; Fridley and Sax, 2014) count as well? Opinions differ on these issues. Webber & Scott (2012) suggest that the alien status of a species should be drawn based on its projected dispersal envelope alone, regardless of whether or not humans have played a role in moving individuals or propagules. Gilroy et al. (2017) and Olenin et al. (2017) generally agree with the point made by Webber & Scott (2012), but rather than disregarding human-mediated dispersal entirely, the former authors distinguish between direct transport by people on the one hand and dispersal that the organisms are able to accomplish on their own strength on the other hand. Specifically, Gilroy et al. (2017) state that their viewpoint affords protected status to those organisms whose only method of dealing with human-induced habitat modification (such as climate change) is to enter previously uninhabited regions. On the other hand, Fridley & Sax (2014) report that the unassisted migration of Red Sea organisms northward through the Suez Canal has caused substantial ecosystem alterations in the Mediterranean Sea, which would never have felt the impacts of these species had the canal never been constructed. Essl et al. (2018) argue that these Red Sea introductions can hardly

be considered natives, as can other taxa that use human-made corridors such as bridges between islands and the mainland.

III. Theory of establishment

III.1: Role of evolutionary history

When introduced into a foreign environment, a species is faced with a range of environmental variables that differ from those in its native range. For instance, one can easily imagine that the alien organism may encounter novel diseases and predators, which it is not evolutionarily equipped to deal with, leave behind preferred food sources, or be forced to exist under suboptimal temperature or other abiotic environmental conditions. It is not surprising, therefore, that only a fraction of NIS develop invasive characteristics (García-Berthou et al., 2005; Kettunen et al., 2009; Faulkner et al., 2014). An often-applied “rule of thumb” that reflects this reality is the so-called ten percent rule, formulated by Williamson and colleagues back in 1996 (Williamson & Fitter, 1996; García-Berthos et al., 2005). This principle states that 10 % of all imported species are introduced into the wild, 10 % of all introduced species establish free-living populations, and 10 % of all established species become invasive (proportions of 5-20 % are also considered as falling within the tens rule) (Williamson & Fitter, 1996; García-Berthos et al., 2005). However, these proportions may be much higher, especially in vertebrates (Jeschke, 2014) and in intentionally introduced organisms (such as crops, biocontrol agents, and economically interesting freshwater fishes; Williamson & Fitter, 1996; García-Berthos et al., 2005).

On the other hand, leaving behind the environment in which they originally evolved, NIS also leave behind various competitors, predators, and parasites that may have adversely affected them in their native range (Boudouresque & Verlaque, 2012; Rejmánek & Simberloff, 2017). The enemy release hypothesis considers this loss of natural enemies to be the main cause of biological invasions, although empirical tests have revealed mixed support for this hypothesis (Jeschke, 2014), indicating that other factors may influence invasion success. At the same time, organisms in the invaded range have never encountered the NIS and thus tend to lack adaptations to deal with the competition, predation, or other antagonistic effects of the new species (Fridley & Sax, 2014). As a result, the NIS gain a large competitive advantage over the native species and become IAS. The two aspects of enemy release and recipient ecosystem naiveté fit well with the predictions of the evolutionary biological invasion framework of Fridley & Sax (2014), which views biological invasions as the result of the mismatch of the invader’s and the native’s evolutionary histories. These authors based their hypothesis on three principles: 1) evolutionary processes, by their very nature, can never lead

to completely perfect adaptations to the environment, 2) the more niches available to, and competition experienced by, an organism, the better adapted an organism becomes, and 3) many locations across the globe share similar environmental parameters. Due to the above points, the evolutionary hypothesis of biological invasions predicts that species from phylogenetically diverse regions have a higher chance of becoming invasive elsewhere than do species from areas of limited diversity. The IAS may be evolutionarily fitter in its new environment than the native species for that environment. Conversely, low-diversity regions should be especially susceptible to invasions. The authors found both predictions to hold true for a range of invasive taxa and invaded ecosystems.

III.2: Role of recipient ecosystem characteristics

Although Fridley & Sax (2014) were original in their attempt to frame a modern hypothesis of biological invasions in evolutionary terms, some of the conclusions that they drew had been circulating in the literature for several decades prior. The idea that biodiversity and habitat integrity increase resistance to biological invasions was first formulated by Charles Elton in 1958 and is known as the biotic resistance hypothesis (Peng et al., 2019). The basic principle behind the biotic resistance hypothesis is that any natural environment harbours a certain number of potential niches filled by species. In species-rich environments, most or all of these niches have been taken up already by native species, while in species-poor environments, many niches are empty and can be usurped by foreign species (Jeschke, 2014). Subsequent examinations of natural communities have, however, yielded mixed support for Elton's hypothesis. In fact, while biodiversity may play a role in protecting against invasions at smaller scales, it might exhibit a positive correlation with susceptibility to invasion at larger spatial and time scales, a phenomenon known as the invasion paradox (Jeschke, 2014; Peng et al., 2019). On the other hand, Peng et al. (2019) conducted a meta-analysis of plant invasions across the world and concluded that little empirical support exists for the invasion paradox.

III.3: Role of anthropogenic factors

Anthropogenic habitat alteration may also give NIS a competitive advantage over natives by redistributing selective forces (Valéry et al., 2008), so that previously inconspicuous non-native taxa may suddenly start to dominate following disturbance (Occhipinti-Ambrogi & Galil, 2010). The role of external human-mediated factors in biological invasions is at the heart of the so-called driver vs. passenger dichotomy, which presents two alternative frameworks for explaining the impact of IAS on their recipient environments. The driver view states that IAS are the cause of decline of native species populations, while the passenger view considers IAS to be facilitated by the same human impacts that originally caused the native species'

decline (Didham et al., 2005). The driver-passenger paradigm can be used to determine how an ecosystem will respond to the removal of a NIS (Bauer, 2012). According to the driver view, removing the IAS will allow the native populations to recover. According to the passenger view, removing the IAS will have minimal impacts on native populations because the habitat is already too degraded to support the natives (Didham et al., 2005). It should be noted that distinctions between drivers and passengers are not always clear-cut: for example, the IAS may have initially been a passenger, but once established may still directly compete with native species (Didham et al., 2005). Such initial passengers, which are facilitated by habitat disturbance but then exert negative influences on native ecosystems, often through positive feedback loops, are termed “back-seat drivers” by Bauer (2012).

The passenger model implies that IAS are often able to benefit from the same anthropogenic disturbance that harms native species. Various studies have shown that those alien species that can establish themselves are typically not a random subset of all imported species, but generally exhibit a suite of traits that correspond to high environmental adaptability (Karatayev et al., 2009; Cardeccia et al., 2018; Novoa et al., 2020). In addition, the very act of translocation imposes another trait-selection filter on species, as those traits that are considered desirable by humans are often conducive to successful establishment (Rejmánek & Simberloff, 2017). Examples include both traits actively selected for by humans and traits that are indirectly advantaged during the transportation process. Rejmánek & Simberloff (2017) list several examples from the literature of deliberate trait selection: ornamental plants, whose often colourful fruits tend to invoke dispersal by birds; pasture grasses, which are bred to “invade” new plots of land; fast-growing trees used in forestry and erosion control; and edible fish that typically reproduce rapidly. Novoa et al. (2020) note that, at least for organisms arriving as “stowaways” or contaminants of cargo, the rigours of transportation ensure that only the hardiest, most tolerant species survive the trip to the new location.

Human-mediated transport not only selects for certain traits, but it also has the capacity to repeatedly introduce individuals so that the founding population is large enough to overcome any Allee effects (Lockwood et al., 2005). In addition, increasing propagule pressure brings with it increasing chance that a reproductive individual will end up in a suitable habitat (Lockwood et al., 2005; Boudouresque & Verlaque, 2012) and, at the level of multispecies introductions, that one of the introduced species will be especially well-adapted to the new environment (a kind of sampling effect; Buckley & Catford, 2016). The propagule pressure hypothesis considers large propagule pressure (defined as both the number of individuals introduced at once and the frequency with which introductions occur) to be the main cause of biological invasions (Occhipinti-Ambrogi, 2007; Boudoresque & Verlaque, 2012; Jeschke, 2014). Unlike many other hypotheses of biological invasions, this particular framework has received a considerable degree of support from numerous studies (Jeschke, 2014).

Researchers of biological invasions increasingly recognise the role of propagule pressure, rather than isolated traits of the invader or the invaded ecosystem, in determining the success of NIS establishment, hence the growing application of this framework in studies since 2000 (Jeschke, 2014). As an example, García-Berthou et al. (2005) studied invasive freshwater fish worldwide and noted that many of the worst IAS they considered (such as grass, silver, and bighead carp) nonetheless have low establishment successes due to stringent reproductive requirements. The fact that these animals still are invasive displays the importance of propagule pressure. Conversely, the authors found that species with the same ecological requirements would sometimes differ greatly in impact and range outside the species' native environments, with the differences generally attributed to increased occurrences of introductions in some areas compared to others.

IV. Marine NIS: impact and modes of introduction

IV.1: Impact of marine NIS worldwide

Invasive species impact all major ecosystems, including the marine realms. In fact, invasive species represent one of the chief threats to marine ecosystems worldwide (Seebens et al., 2013). Especially the Mediterranean Sea, North Sea, Hawaiian Islands, and western coast of the United States suffer greatly from IAS (Figure 1; Occhipinti-Ambrogi, 2007; Molnar et al., 2008), while polar regions are currently not greatly impacted by marine bio-invasions but may become increasingly invaded in the future as a result of e.g. climate change and the resulting increase in shipping through these regions (Molnar et al., 2008; Cheung et al., 2009; Brown et al., 2016). Although the danger marine IAS currently pose to seas and oceans worldwide is generally acknowledged within the literature, it appears that marine systems tend to suffer less from biological invasions than terrestrial or freshwater systems do (Anton et al., 2019), and several hypotheses have been put forth to explain this observation. One possible reason is that marine habitats are less constrained by geographical boundaries, and hence more connected, than terrestrial or freshwater habitats, thereby allowing native taxa to escape competition from IAS by migrating rather than succumbing to extinction (Cheung et al., 2009; Anton et al., 2019). In addition, the high dispersal capacity of

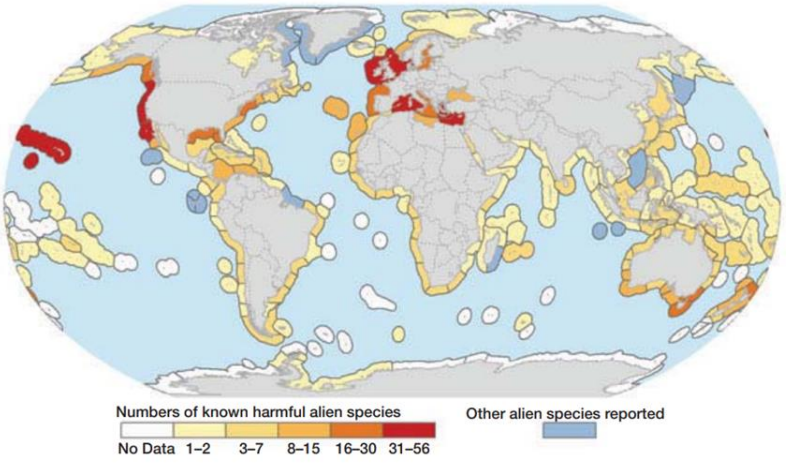

Figure 1: Numbers of recorded marine IAS across the world. Source: Molnar et al., 2008.

many marine taxa may further facilitate this temporary escape from IAS, preventing the foreign species from completely extirpating the natives, which are then able to recolonize their original area at a later date (Anton et al., 2019). However, it may also be that the perceived resistance of marine environments is caused by suboptimal choices of endpoints and the grouping together of species with highly dissimilar impacts, as illustrated by Thomsen (2020). As explained previously, invasive species typically share a set of traits that enable adaptation to and utilization of a wide range of resources, including artificially provided ones, in the host community. Cardeccia et al. (2018) demonstrated a consistency of traits among the 68 most widely established NIS in various European marine waters, noting that most of these species exhibited a benthic lifestyle and an iteroparous reproductive cycle with at least one pelagic life stage. Species with a generalized feeding strategy, such as suspension feeders and opportunistic predators, also showed increased tendency to establish outside their native ranges. Dafforn et al. (2009) note that many marine invasive species are able to rapidly colonize previously sterile surfaces, and as such are benefitted by the introduction of man-made structures such as pontoons, marinas, and buoys.

IV.2: Introduction via ballast water

The largest source of marine NIS is ballast water from ships on trans-oceanic voyages (Molnar et al., 2008; Seebens et al., 2013; Casas-Monroy et al., 2015). The chance of a new introduction from one community to the other via ballast water increases with community dissimilarity (Seebens et al., 2013) and hence with increasing distance between the donor and recipient communities (Casas-Monroy et al., 2015). Conversely, invasion success decreases with growing environmental mismatch, especially in terms of temperature and salinity, between donor and recipient communities (as can be inferred from Figure 2; Seebens et al., 2013; Casas-Monroy et al., 2015). As a result, it is not only the shipping intensity a given area receives but also the environmental variables of that area that determine how susceptible it is to new bio-invasions. The North Sea, for example, receives a large amount of shipping traffic yearly but is comparatively at less risk of new bioinvasions than other seas due to the relatively low environmental match with other major shipping ports (Seebens et al., 2013). (Note that the preceding statement deals with predicted future invasions, as opposed to the impact that the

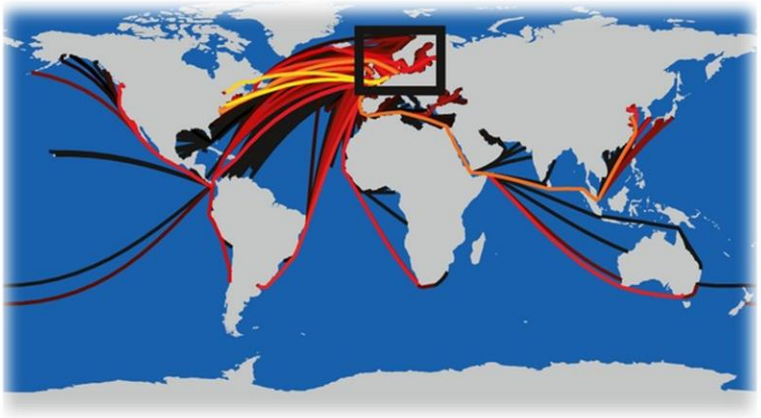

Figure 2: Estimated risks of various shipping pathways of introducing invasive species to the Northern European seas (black box). Lighter colours correspond to higher-risk pathways. Source: Seebens et al., 2013.

many already established IAS have on the North Sea.) As global distance generally correlates with decreased similarity in both biotic and abiotic conditions, it can be expected that the length of the shipping route followed by a NIS is a good determinant of introduction success. Indeed, records indicate that most NIS invasions occur over intermediate distances of 8000 - 10000 km (Seebens et al., 2013). Short distances give a low probability of introducing a truly non-native species, while long distances are typically covered less frequently in maritime traffic and also bring along a decreased chance for the NIS to survive the voyage (Seebens et al., 2013). The probability of introduction of a NIS is, therefore, also a function of survival during transport, which is linked to travel time. Furthermore, the amount of ballast water carried by the vessel is also a determinant of a ship's probability of introducing NIS (Seebens et al., 2013).

To combat this major source of marine NIS and IAS introductions, measures to treat ballast water have been implemented to various degrees worldwide. Ballast water exchange is the most common preventive measure because it is comparatively less expensive and easier to carry out by ship owners than other preventive measures are (Balaji et al., 2014). Treatment entails the replacement of some or all ballast water mid-ocean, rather than at the port of destination as has commonly been the practice (Casas-Monroy et al., 2015). The rationale behind a mid-ocean replacement is that organisms from the port of origin are released not at a new port, which might prove inhabitable by them, but in the open ocean, where they are unlikely to survive. Conversely, any planktonic species taken up with the new ballast water in the open ocean are unlikely to find a suitable environment in a coastal setting (but see Casas-Monroy et al., 2015). Ballast water exchange is most effective for vessels undertaking larger voyages (> 200 nautical miles from the shore) rather than staying within a coastal region, and for vessels making a transition between saltwater and freshwater, where the osmotic shock generated kills the NIS in the ballast water upon arrival (Casas-Monroy et al., 2015). Other methods of ballast water treatment include electrochlorination, ultraviolet irradiation, filtration, deoxygenation, and cavitation, often implemented in various combinations (Balaji et al., 2014). The efficiency of these various methods of treatment in preventing new NIS establishments is variable. On one hand, Seebens et al. (2013) demonstrated that even moderate efforts to treat ballast water can still lead to a large reduction in invasion probability, due to the multiplicative effects of ballast water treatment at each successive port during a voyage. Treating 25% or 50% of the ballast water at each port would result in an overall reduction in invasion probability of 56% and 82%, respectively, according to the model developed by the authors. On the other hand, it has been shown that up to 30% of all planktonic organisms in a ballast water tank remain after ballast water exchange (Casas-Monroy et al., 2015), and that ballast water exchange alone is not suitable for achieving modern regulatory standards concerning ballast water biosafety (Balaji et al., 2014).

Legislation mandating ballast water exchange has historically not always been implemented and enforced to the same degree everywhere (Sardain et al., 2019). As of 2018, Ojaveer et al. (2018) report that only two global instruments regarding NIS control were legally binding: the United Nations Convention on the Law of the Sea (UNCLOS) (which requires all member states to take the necessary measures to prevent and control NIS) and the International Convention for the Control and Management of Ships' Ballast Water and Sediments (BWMC) (which establishes maximum limits for the amount of live organisms and propagules that are tolerated per unit of ballast water). The BWMC's more stringent regulations compared to earlier legislation meant that many ships had to supplement or replace their ballast-water exchange practices with some of the other methods listed above (Balaji et al., 2014).

IV.3: Introduction via biofouling

Together with unintentional transport via ballast water, biofouling of ships and other mobile structures is one of the major vectors for marine IAS (Molnar et al., 2008). Especially those areas of the vessel that are more difficult to clean or are more easily overlooked (such as the rudder, propellers, and intake pipes) can carry dense epifaunal communities over long periods of time (Lane et al., 2018). Again, global shipping networks are largely responsible for this introduction via biofouling, although Lane et al. (2018) showed that recreational vessels can be considerable sources of biofouling NIS as well, due to the propensities of these vessels for low travel speed, irregular maintenance schedules, and visitation of pristine marine habitats that are especially sensitive. Biofouling has traditionally been controlled through the application of toxic paints on underwater surfaces of ships and other vessels, but this practice has been more strictly regulated since 2008 after tributyltins – the active components of many of these paints – were shown to cause mortality in cultured oysters and to adversely affect other marine organisms (Ojaveer et al., 2018). Hence, in recent years, a trend towards development and application of anti-fouling paints that are less hazardous for the marine environment can be noted (Hopkins et al., 2016). One drawback of paints in general is that they cannot be applied to all surfaces that may harbour NIS communities, such as fishing nets (Hopkins et al., 2016). Another widely used alternative to anti-fouling paints is manual removal of the communities, for instance via divers, although this method is particularly labour-intensive (Atalah et al., 2016). Due to the limitations of anti-fouling paints and manual removal methods, alternative methods to combat biofouling are under development (Joyce et al., 2019), such as application of steam or hot water (Joyce et al., 2019), desiccation (Hopkins et al., 2016), and the use of indigenous predators of biofouling communities (Atalah et al., 2014; Atalah et al., 2016).

IV.4: Introductions via intentional importation of species

Aquaculture presents a third major gateway for marine NIS introduction (Naylor et al., 2001; Weigle et al., 2005; Molnar et al., 2008). Naylor et al. (2005) note that escapes from fish farms are almost inevitable, a situation that has resulted in farmed salmon and other commercially important species establishing themselves outside their native range in various parts of the world. The escapees then compete for food and spawning areas with other species as well as with genetically distinct, wild conspecifics, often threatening the latter through hybridization. Sedentary organisms, such as molluscs, also pose an escape risk if they begin to release their planktonic larval stages into the water; for instance, the oyster *Magallana gigas* was introduced into the North Sea via this manner (Faust et al., 2017). Even if the aquaculture facility is completely escape-proof, NIS introduction can still occur when shipments of aquaculture livestock contain stowaway organisms (such as parasites) hiding in the transport containers or on the livestock itself (Weigle et al., 2005; Grosholz et al., 2015). Although not as large a vector of NIS spread as biofouling or ballast water, aquaculture does result in more directed transport of NIS, as non-native species employed at a given location have most likely been chosen by humans to thrive within the environmental conditions of that location, as well as to exhibit various features such fast growth that could translate into invasiveness later on (Weigle et al., 2005; Grosholz et al., 2015).

The trade in marine organisms for non-aquaculture purposes, such as for fishing bait or aquarium use, also presents a substantial pathway for NIS introduction (Weigle et al., 2005). The widely publicized spread of the alga *Caulerpa taxifolia* throughout the Mediterranean Sea occurred after its accidental release via wastewater from a public aquarium (Montefalcone et al., 2015). Similarly, various species of lionfish (*Pterois* spp.) became established in the Western Atlantic Ocean after their release from captivity, possibly after outgrowing the tanks they were kept in, and are now adversely impacting native marine communities through predation and non-consumptive effects (Lyons et al., 2020).

The dispersion of NIS via aquaculture and other forms of trade in organisms is most effectively reduced through stricter and more enforced legislation, as well as through increasing awareness of the consumers, middlemen and producers in these trades (Weigle et al., 2005; Grosholz et al., 2015).

IV.5: Introductions via canals

Canals between formerly isolated bodies of water can also allow marine species to expand their distributions into areas well beyond their natural dispersal ranges. Many such migrants have been documented to adversely affect the species originally inhabiting the colonized area (Molnar et al., 2008). A well-known example is the massive colonization of the Mediterranean

Sea by Red Sea species via the Suez Canal (Occhipinti-Ambrogi & Galil, 2010; Guy-Haim et al., 2017; Lyons et al., 2020). In this case, many of the new species have been shown to compete with and/or predate on natives, sometimes displacing the latter locally.

V. Detection, monitoring and management of marine NIS and IAS: current strategies and recent technological advances

V.1. The necessity of prioritization

Marine IAS are almost always impossible to completely eradicate once they have been sufficiently established in their new locations to form self-sustaining populations (Bott, 2015; Ojaveer et al., 2018). Cases of successful eradication of established marine IAS populations do exist, but they typically involve highly isolated locations that were easily separated from the main seas for treatment and/or pests that were detected very early, before they began to spread (Bott, 2015). For most marine NIS, management of already present populations is generally regarded as unrealistic (Novoa et al., 2020). Therefore, successful programs tend to focus on preventing the introduction of NIS in the first place, and only if this fails, removing or containing the organisms before they begin to exhibit deleterious effects on the recipient ecosystem (i.e., mitigation) (Faulkner et al., 2014; Trebitz et al., 2017; Ojaveer et al., 2018). In both cases (prevention and mitigation), costly monitoring and remediation schemes must usually be employed; hence, an effective prioritization scheme of which species and pathways to target first is essential (Campbell et al., 2007; McGeoch et al., 2016).

V.1.1: Prioritization of species

Effective prioritization considers three key factors: species, pathways, and sites (McGeoch et al., 2016; Novoa et al., 2020). Species of concern (i.e., IAS or species with the potential to be so) are typically included into so-called watch lists, which are constructed based on the degree of similarity between a species' native and introduced range, as well as on the past invasive behaviour of the species outside its native range (Bishop & Hutchings, 2011; Trebitz et al., 2017). Together with propagule pressure (i.e., the number of reproductive units released per introduction event, which is dependent on human behaviour), these two features are widely considered to be important predictors of invasiveness (Faulkner et al., 2014). Watch lists are increasingly used in marine IAS and NIS management (Bishop & Hutchings, 2011) due to the fact that they are comparatively easy to construct and implement, at least compared to detailed species-specific risk assessments (Faulkner et al., 2014). One example, relevant to this study, is the Harmonia+ protocol, which is used in Belgium to assess whether NIS present a threat to native ecosystems (D'hondt et al., 2015), and which has resulted in the creation of a nation-

wide watch list of 38 species (Belgian Biodiversity Platform, 2020). None of these organisms are marine; this likely reflects the general lack of attention given to marine IAS compared to terrestrial and freshwater NIS (Ojaveer et al., 2018). However, the Harmonia+ protocol can also be applied to marine NIS and may be used as such in the future (D'hondt et al., 2015; McGeoch et al., 2016). In addition to being useful across taxa and environments, it is also constructed according to well-established protocols (McGeoch et al., 2016), and as such presents an important contribution to the construction of watch lists, which tend to be set up in a non-standardised manner (Faulkner et al., 2014).

Popular as they are in marine NIS and IAS management, watch lists have important shortcomings. Bishop & Hutchings (2011) mention three important issues that arise when surveys of at-risk areas are conducted focusing on a predetermined list of species. First, watch lists ignore unique (a)biotic factors of the recipient ecosystem relative to those of ecosystems where the species were known to be invasive. Consequently, the impact of the species on the environment in question is difficult to predict from impacts (or the lack thereof) on other areas of the world. Second, watch-list based surveys do not include native species and (often) currently non-invasive NIS, and hence cannot establish baselines against which to compare future observations. Third, watch lists may lead to the overlooking of new NIS which, although not currently recognized as a threat, may become invasive later. In addition, as noted by McGeoch et al. (2016), using a species-based prioritization scheme is likely to be unsuccessful when species are introduced unintentionally, as are most marine NIS (Ojaveer et al., 2018). In this case, better results can be achieved by focusing on invasion pathways, taking into account both the means of spreading and the propagule loads for each pathway (McGeoch et al., 2016). Monitoring schemes that employ such a pathway-focused approach are referred to as broad-spectrum in comparison to watch-list-based or target-based surveys (Trebitz et al., 2017).

V.1.2: Prioritization of pathways

Invasion pathway analysis involves the compilation of all the pathways known to introduce non-indigenous and/or invasive species into a region in the past, and then prioritizing these pathways according to both their frequency of occurrence (such as intensity of shipping) and the amount of propagules released each time such a pathway is followed (such as number of organisms released in one quantity of ballast water) (McGeoch et al., 2016; Novoa et al., 2020). The identification of relevant vectors of introduction subsequently enables more effective legislation to be written and implemented. Much global legislation has, in fact, already been drafted targeting invasion pathways, in the recognition that they represent key aspects in facilitating biological invasions (Hulme et al., 2008; Hulme, 2009). For example, the growing research on the role of ballast water as a major introduction pathway for various marine NIS

has recently led to the introduction of the international D-2 standards for ballast water management, which require an increase in protective treatment management relative to previous legislation (Casas-Monroy et al., 2015). Especially for the European Union, documentation of biological invasions has been quite extensive and hence has resulted in a large amount of regulatory literature, although this is still seldom implemented effectively (Hulme et al., 2008). However, finding common ground between these regulations has proven difficult, as pathways tend to be delineated and defined rather vaguely, with potential overlap of definitions depending on the source (Essl et al., 2015). As an illustration, Hulme et al. (2008) record that, at the time of their writing, the Global Invasive Species Database distinguished almost thirty different pathways, while the Convention on Biological Diversity recognized only two. As with the definitions of non-native and invasive, no international consensus exists on how pathways must be delineated, leaving pathway-based prevention and mitigation measures largely at the hands of national and local authorities, even though NIS and IAS are a problem transcending political boundaries (Gilroy et al., 2017; Hulme et al., 2008; Hulme, 2009).

Invasion pathways are not static in time. They may be altered as a result of changing political and economic ties among countries, or due to changes in global and regional consumer demand, and hence the relative importance of any given pathway may increase or decrease over the years (Essl et al., 2015). Nor do different pathways work independently of each other. Generation of an effective risk map for biological introductions and invasions requires a model that links the various pathways via nodes (such as ports or other shipping stations) and has information on the relative importance of each pathway and node (Hulme, 2009).

V.1.3: Prioritization of sites

The third key factor identified by McGeoch et al. (2016) for effective prioritization measures is the recipient ecosystem. Propagule pressure may have little effect if the environment is too hostile to the NIS, such as if competition with native species is very intense (Seebens et al., 2013; Hulme, 2009; Novoa et al., 2020). Local abiotic factors, of course, play an important role as well. A widely acknowledged fact is that NIS originating in regions showing high environmental disparity with that of the host ecosystem have low chances of successfully establishing, and hence represent a lesser threat (Seebens et al., 2013; Fridley & Sax, 2014; Essl et al., 2015; Novoa et al., 2020). For marine NIS, temperature and salinity are the most important abiotic determinants of establishment success (Sardain et al., 2019). Locations that exhibit a high chance of having NIS introduced into them (susceptible sites) and those that are especially vulnerable to introductions (sensitive sites) must receive comparatively greater priority than other areas (McGeoch et al., 2016). The susceptibility of a site is determined by both human and natural factors, for example, the intensity of shipping received on the one

hand and the climatic suitability for particular NIS on the other. Sensitivity is largely dictated by a region's biodiversity, and refers to the impact that a biological invasion, regardless of its actual chance of occurring, would have on the ecosystem. Examples of sensitive sites would be regions harbouring rare natives and regions delivering especially important ecosystem services (McGeoch et al., 2016). Even if preventive measures have failed to intercept NIS settlement, knowledge of the environment may still be able to slow down the spread of these species. For example, natural barriers may prevent the range expansion of NIS following establishment of viable populations (Hulme, 2009). Knowledge of such boundaries is essential for prioritization, as it is especially the breaks in these barriers that are prone to allowing spread of non-indigenous and invasive species (Hulme, 2009). Linear natural or anthropogenic structures, such as railways or rivers, are important dispersal corridors for NIS, but are less important when connecting heavily human-modified environments, relative to when they allow NIS to access relatively undisturbed natural environments (i.e., sensitive sites) (Hulme, 2009).

V.2: Issues with traditional NIS survey approaches

Early detection of potential or already established marine NIS, before they become invasive, is typically a labour-intensive enterprise. The amount of effort that must be invested is ultimately determined by the trade-off between easier detection and more difficult eradication as the NIS population increases (Trebitz et al., 2017). Traditionally, detection and quantification of NIS in marine areas has involved morphology-based methods, which rely on visual identification of the organisms (Duarte et al., 2021). A morphology-based survey requires collection of as many samples and sample types as possible, in order to increase the chance of detecting rare or as yet non-abundant taxa (Campbell et al., 2007). Typical sampling protocols reported by Campbell et al. (2007) make use of scrape samples, sediment core samples, crustacean traps, settlement plates, nets, and visual inspections by trained divers, among others. In their overview of sampling protocols, these authors also emphasize the importance of site choice for conducting a representative survey. Different protocols use different criteria for choosing sites, generally trying to strike a balance between taxonomic rigor and facility of use. For example, the most widely used protocol (Hewitt & Martin protocol; Hewitt & Martin, 2001) utilizes a Poisson model to determine which sites to target in order to maximize the number of rare species located. In contrast, so-called rapid assessment methods target sampling at locations that can be easily reached from the shore or wharf. The lack of coverage depth of these methods is then counteracted by their increased facility of use. Campbell et al. (2007) also note that morphological sampling protocols can be constructed to target specific types of NIS; for example, the Chilean government's standardized survey for detecting abalone aquaculture escapees, or the use of settlement plates for obtaining information on the epibenthic fouling community in a particular area.

Morphology-based sampling protocols do have important shortcomings that may hamper their usefulness in proactive NIS detection. Difficulties of morphology-based detection include the high amount of labour required to conduct adequate surveys, the global lack of taxonomic expertise for accurate identification of all specimens, and the general inability to detect certain species (von Ammon et al., 2018; Duarte et al., 2021). Small species with cryptic habits, such as many members of the meiofauna, can be especially difficult to differentiate from each other using visual methods alone (Tang et al., 2012; von Ammon et al., 2018; Duarte et al., 2021). The same goes for many juvenile organisms, which often represent the very life stages that are most adept at spreading (Ojaveer et al., 2018). Relevant taxonomic expertise is declining worldwide (Bott, 2015) and hence identifications must often be conducted by informally trained researchers rather than specialists (Campbell et al., 2007). This fact is cause for concern, as the incorrect assignment of species' origin may have important consequences on the results and predictions of a NIS survey (Bishop & Hutchings, 2011; Zaiko et al., 2015). Destructive sampling methods, such as trawling the sea floor or scraping submerged surfaces, may also risk harming native populations (Campbell et al., 2007; Fraija-Fernández et al., 2020). Many sampling techniques are also overly selective and do not give an accurate representation of either the native or the introduced community at a given area (Dafforn et al., 2009; Tait et al., 2018; Fraija-Fernández et al., 2020). Regarding settlement plates, which are widely used in marine NIS monitoring (Campbell et al., 2007), Dafforn et al. (2009) noted that site and orientation of these plates already had an impact on the communities that settled on these structures. Tait et al. (2018) conclude that settlement plates are generally inept at detecting rare species, with over eighty separate plates needed to consistently detect rare species. For this and other morphology-based methods, reliable species detection typically occurs only if the density of the target species is already high, a fact that makes preventive detection of NIS (i.e., before they become highly abundant) difficult (Brown et al., 2016).

V.3: The promise of DNA-based techniques

V.3.1: Overview of DNA-based monitoring techniques

To combat these shortcomings of morphology-based methods, molecular approaches to NIS surveillance are increasingly employed worldwide (Bott, 2015; van der Loos & Nijland, 2020). These DNA-based methods do not require detailed taxonomic knowledge and allow detection of rare or cryptic organisms that might be missed in morphological samples (Zaiko et al., 2015). These methods discriminate between species or higher groupings using barcodes, or short, taxon-specific segments taken from the same area of the genome across multiple specimens, and compare the detected barcode sequences against those recorded in a reference database (Kress et al., 2015). Either the organisms are individually identified using

barcodes, or a sample of the entire community is taken and identified using these barcodes without first isolating individual morphospecies. The first approach is referred to as DNA barcoding, while the second approach is referred to as DNA metabarcoding (Zaiko et al., 2015). DNA can be extracted for (meta)barcoding from collected specimens, but it can also be obtained directly from the environment without having to obtain the organisms themselves (Taberlet et al., 2012; Kress et al., 2015; Holman et al., 2019). DNA is known to exist in detectable concentrations in the environment outside of living tissues, as a result of processes such as excretion, loss of cells, and decay of dead organisms, in the form of so-called environmental DNA or eDNA (Creer et al., 2016; Holman et al., 2019). Therefore, a sample of sediment, water or even air has the potential to give valuable information about the organisms living in the vicinity, even if they cannot be located themselves (Trebitz et al., 2017).

V.3.2: Overview of common DNA barcodes

The barcode of choice to be used for DNA-based monitoring of marine NIS, and indeed for any other species detection program, depends on the taxon screened for. A large proportion of research on metazoan organisms distinguishes species or higher groupings using a fragment of the mitochondrial cytochrome oxidase subunit 1 (COI) gene as barcode (van der Loos & Nijland, 2020). This fragment has traditionally been a fixed sequence of 658 base pairs (bp) known as the Folmer fragment (Bucklin et al., 2011; van der Loos & Nijland, 2020). Recently, though, a shorter COI barcode of 313 bp (the so-called Leray fragment) has been increasingly employed due to its increased efficiency and amenability to the popular Illumina MiSeq sequencing platform (van der Loos & Nijland, 2020). The COI gene is generally considered a universal genetic marker for animal species (Valentini et al., 2009; Taberlet et al., 2012) due to its comparatively high interspecific and low intraspecific genetic distances (Bucklin et al., 2011). The variation exhibited at the species level does not overlap with the variation found at higher taxonomic levels: this pattern leads to a so-called “barcoding gap”, which is the main reason the COI gene is the most popular barcode location (Bucklin et al., 2011). The high rate of interspecific divergence exhibited by this gene is due to the fact that it has a high rate of evolution (von Ammon et al., 2018). The COI gene has been used to successfully identify marine animals from many different phyla to the species level, although some groups—such as anthozoans and copepods—have seen less success using this method, either because their rates of mtDNA evolution are unusually low or because the primers used to amplify the barcodes show selective annealing (Bucklin et al., 2011; Zaiko et al., 2015). COI barcodes have also proven ineffective for plants and fungi, again due to lower rates of mtDNA evolution in these kingdoms (Kress et al., 2015).

Van der Loos & Nijland (2020), in their review of marine metabarcoding studies, found that the mitochondrial 18S rRNA gene is as commonly used as a marker as the COI gene. The 18S

rRNA gene is part of the nuclear DNA that codes for the small subunit of the ribosomes (von Ammon et al., 2018). Its primer-binding sequences are more conserved than those of the COI gene, making primer design easier (van der Loos & Nijland, 2020) and allowing its application for a broader range of taxa than the COI gene (Creer et al., 2016). Another result of the greater inter-taxon similarity exhibited by the 18S rRNA gene is that it exhibits higher rates of successful amplification (Duarte et al., 2021). However, the lower variability in 18S rRNA sequences found across phylogenetic groupings also results in lower species identification ability and a tendency to underestimate true species diversity (Holman et al., 2019; van der Loos & Nijland, 2020; Duarte et al., 2021), although von Ammon et al. (2018) found that the 18S rRNA gene was more accurate at identifying species from mixed fouling community samples than the COI gene was. Being part of the nuclear DNA rather than the mtDNA, copies of the 18S rRNA gene are much fewer in number inside any cell than the COI gene copies are, and hence detection of rare species is more difficult with the former marker (Bucklin et al., 2011; Duarte et al., 2021).

Besides the 18S rRNA gene, other ribosomal subunit genes are also typically used for barcoding and metabarcoding, such as the 12S and 16S rRNA genes (van der Loos & Nijland, 2020). These two genes, unlike the 18S gene, are part of the mtDNA, and thus exist in large copy numbers inside the cell just as COI does (van der Loos & Nijland, 2020). It can thus be expected that the discriminatory abilities of these barcode regions are quite high, and indeed, the 12S gene region shows especially high resolution for teleosts (Fraija-Fernández et al., 2020; van der Loos & Nijland, 2020). However, their sequences are generally more conserved across multiple taxa than those of COI, and they exhibit frequent indels, two conditions that limit the use of the 16S and 12S rRNA regions (Bucklin et al., 2011; van der Loos & Nijland, 2020). In addition, reference sequence databases for these two rRNA genes are currently limited in scope, and many species are lacking representative sequences in these databases (Shaw et al., 2016; van der Loos & Nijland, 2020). Obviously, each different DNA barcode has its own strengths and weaknesses, and the choice of marker is heavily dictated by a trade-off between specificity and taxonomic coverage (Cristescu & Hebert, 2018; van der Loos & Nijland, 2020). It is generally acknowledged that using a combination of markers, for example COI in combination with 18S, increases the ability of a study to accurately detect different species relative to when only one marker is used (Creer et al., 2016; Cristescu & Hebert, 2018; van der Loos & Nijland, 2020).

V.3.3: DNA sequencing approaches

The implementation of DNA barcoding and metabarcoding for ecological studies has been greatly facilitated by the advent of more modern DNA sequencing methodologies (Bucklin et al., 2011; Taberlet et al., 2012). Referred to as next-generation or high-throughput sequencing

(NGS or HTS, respectively), these methods allow for many sequencing reactions to be run in parallel in a short amount of time (Valentini et al., 2009). NGS and HTS methods offer numerous advantages over the previously used workflow, which relied on capillary electrophoresis (Valentini et al., 2009). Because the modern high-throughput methods allow a greater depth of sequencing (i.e., allow many more reads per gene region, in this case the barcode), they are more apt at detecting rare variations in the barcode within the sample and therefore have a greater detection power (Valentini et al., 2009; van der Loos & Nijland, 2020). Greater sequence depth can also offset biases due to PCR amplification and primer bias (van der Loos & Nijland, 2020). In addition, financial and monetary costs are substantially reduced in the HTS and NGS methods by the removal of the previously necessary step of cloning all PCR-amplified DNA fragments into bacterial hosts before sequencing (Valentini et al., 2009).

Figure 3: Oxford Nanopore Technologies MinION sequencer connected to the USB port of a laptop. Credit: Oxford Nanopore Technologies (https://nanoporetech.com/about-us/for-the-media#image123634281&modal=image123634281).

The nanopore method represents a further advancement in sequencing technology. In this method, a single-stranded DNA molecule, driven by an electrical gradient and aided by an enzyme, passes through a pore in a protein membrane. In doing so, the strand partially blocks the current flow across the membrane, with different nucleotides reducing the current to different extents. In this way, the instantaneous current across the membrane indicates which nucleotide is in the pore at that moment, and a record of these instantaneous currents through time determines the sequence of the DNA strand (Deamer et al., 2016). Unlike second-generation sequencing methods (such as Illumina sequencing), nanopore methods are not dependent on synthesis or amplification of DNA, so that they can produce reads of theoretically unlimited length (Kono & Arakawa, 2019). Long read length, in turn, increases species resolution (Brown et al., 2015). Since no imaging equipment or analysis is needed to identify the nucleotides, the sequencing devices can have a small size, short run time, and reasonably low cost (Kono & Arakawa, 2019). A noteworthy example of a nanopore sequencer is the MinION, produced by Oxford Nanopore Technologies from the UK (Figure 3). Unlike the typical second-generation sequencers, the MinION is small enough to be held in one's hand and can be powered by a laptop USB port (Kono & Arakawa, 2019). Loit et al. (2019) reported that the entire MinION-based workflow, from sample collection to interpretation of the results, could be performed in as little as two hours and a half. Due to these advantages, the MinION may be

especially suited to rapid detection of NIS in harbours, without the long waiting times associated with library preparation and sequencing via second-generation methods (Loit et al., 2019).

V.3.4: Concerns with DNA metabarcoding

Despite being a highly sensitive and accurate approach for the detection of marine NIS (von Ammon et al., 2018), metabarcoding has to overcome several barriers that limit its application. The inherent differences in barcodes discussed above, and the resulting trade-off between taxonomic coverage and taxonomic specificity, is one obvious difficulty. Other difficulties arise from the steps bridging sample acquisition and sequencing. Sample processing is an inherent source of bias: for example, sieving and decantation as used to isolate microscopic metazoans tend to lead to the loss of viruses or unicellular organisms from the sample (Creer et al., 2016). The current dependence of most metabarcoding methodologies on pre-sequencing PCR also represents a potential source of errors (Kress et al., 2015; Taberlet et al., 2012). Due to multiple causes, such as differences in primer affinity or the presence of other biological material, DNA is not always representatively amplified from a sample, so that some species may be underrepresented or even absent in the final analysis (Trebitz et al., 2017). Even when DNA extraction and sequencing have proceeded successfully, publicly available databases typically do not contain sequences from all members of a given taxon and/or show a bias towards inclusion of certain groups over others (Kress et al., 2015; Zaiko et al., 2015; von Ammon et al., 2018). Databases may also contain errors and incorrect sequence-species matches (Valentini et al., 2009). Some databases, such as the COI-based Barcode of Life Database (BOLD), also tend to give high false positive rates (Cristescu & Hebert, 2018). For these reasons, compilation of small-scale reference sequences specific to the fauna of a particular study area is often recommended to facilitate future metabarcoding studies in this region (Zaiko et al., 2015). Even the matching of sequence reads to databases can introduce errors, such as when artefactual reads are taken to represent rare species (Cristescu & Hebert, 2018). However, it must be remembered that low species resolution and faulty identifications may also reflect natural causes (such as hybridization) or the flexible interpretation of the species concept itself, and not necessarily methodology-based flaws (Valentini et al., 2009; Kress et al., 2015). Considering all these current difficulties, it is perhaps logical that both Zaiko et al. (2015) and von Ammon et al. (2018) found that a combination of metabarcoding and morphological identification was more informative in marine NIS detection studies than either method by itself. Trebitz et al. (2017) agree that the synchronous use of both traditional and metabarcoding methods currently represents the best strategy for effective NIS detection.

VI. Introduction to the GEANS project

The North Sea is one of the most invaded marine ecoregions in the world (Molnar et al., 2008). Verleye et al. (2020) report 79 established NIS in the Belgian part of the North Sea alone. Numbers of NIS introduced are only expected to continue rising in the North Sea region, as they are for other regions in the world (Sardain et al., 2019). Therefore, effective monitoring programs are needed to enable effective NIS-related legislation to be implemented for this sea. Those parts of the North Sea that fall within the jurisdiction of European Union member countries are covered by the extensive EU environmental legislation, such as the Habitats Directive (Ojaveer et al., 2018) and the Marine Strategy Framework Directive, which was recently succeeded by the EU Biodiversity Strategy for 2030 (European Commission, 2020). Although not directly focused on NIS and IAS, these and many other Union-wide conventions aim for the general improvement of natural areas and as such take measures to control the introduction of non-native species (Ojaveer et al., 2018). Since 2015, legislation aimed specifically at IAS (although not exclusively for marine species) has been in effect within the EU (Tollington et al., 2017). This legislation adopts a black-list approach, where an inter-Union group of experts periodically evaluate potential risk species using fourteen minimum standards (Tollington et al., 2017). The North Sea also receives protection by various non-EU bodies, such as OSPAR (Walday & Kroglund, 2002). OSPAR is a convention by various governments (both within and outside the EU) to follow a modified version of the 1972 Oslo Convention and 1974 Paris Convention to protect the area of the Northeast Atlantic Ocean, including the North Sea (OSPAR, 2020). Together with the Helsinki Commission (HELCOM), a similar intragovernmental agreement aiming for the protection of the Baltic Sea (https://helcom.fi/about-us/), OSPAR has developed a survey protocol for NIS in European ports, with the aim of granting ships exemptions to international ballast water management requirements (as determined by the International Maritime Organization) if it can be demonstrated that the trajectory travelled by a ship does not have a significant risk of picking up or spreading NIS of concern (OSPAR, 2019).

The OSPAR-HELCOM sampling procedure (OSPAR, 2019) is partly based on the Hewitt-Martin protocol and thus focuses on so-called high priority sites within a port, such as active berths and slipways. However, elements of rapid assessment protocols were also incorporated to facilitate implementation of the survey procedure. Each port should have at least three sampling sites, with at least three replicates per site. The port must be sampled twice, once during the spring bloom and once during the summer maximum. At each site, settlement plates are introduced during the spring bloom and retrieved during the summer maximum, while plankton samples are taken per site at both time points. Epifaunal, benthic and fouling communities must be recorded whenever they are most speciose, with traps used to collect mobile epibenthos and sediment grab samples for collection of sediment-dwelling

organisms. In addition, each sampling site requires measurement on physical characteristics (which should include at least both temperature and salinity) and a water sample to be tested for pathogenic bacteria. For each survey, species effort/accumulation curves should be constructed to demonstrate that the sampling effort was sufficient. Although all organisms must be reported, a specific list of target species is used when conducting the risk evaluation for the exemption grant.

This sampling protocol is, however, largely morphology-based (https://northsearegion.eu/geans/) and as such carries all the impracticalities of such NIS surveys, as described previously. A recent EU Interreg project has attempted to streamline the NIS detection and monitoring process in the North Sea using genetic survey methods and is named GEANS (Genetic Tools for Ecosystem Health Assessment in the North Sea) (https://northsearegion.eu/geans/). GEANS is funded by the EU's Interreg North Sea Region program (https://northsearegion.eu) and involves partners from nine different institutions (including VLIZ, the Flanders Marine Institute) located in seven countries (Belgium, United Kingdom, the Netherlands, Germany, Denmark, Sweden, and Norway) (https://northsearegion.eu/geans/). The GEANS project (described as below by the GEANS website, https://northsearegion.eu/geans/) aims to standardize genetic methods for surveying the North Sea biodiversity, with the goal of improving related environmental policy decision-making. In order to accomplish interregional standardization, GEANS is conducting various pilot studies that deal both with direct human impacts on the marine environment (renewable energy constructions, dredging activities, and aquaculture) and with the monitoring of NIS in North Sea harbours. The NIS surveys are performed based on the OSPAR-HELCOM protocol described above, with the difference that metabarcoding will be used to identify all the samples. In addition, the samples will also be identified morphologically. This approach will enable comparison between DNA-based and traditional survey methods on the one hand, and it will also allow past taxonomic knowledge to be integrated into the NIS monitoring continuum on the other hand. Although no routine genetic monitoring occurs within the North Sea area yet, the standardization provided by GEANS will enable cost- and time-efficient regular monitoring in the future (https://northsearegion.eu/geans/).

Objectives

This thesis, which is part of a GEANS pilot study, aims to evaluate the performance of metabarcoding in detecting NIS from both benthic and planktonic communities in the harbour of Ostend (Belgium). To this end, presences and absences of species as revealed by different detection methods are compared.

Benthic communities are sampled by deploying settlement plates at three different locations in the harbour for two months. Afterwards, the communities established on these plates are examined morphologically, and as many taxa as possible are identified. Following morphological examination, the settlement plate samples are metabarcoded with the V4-V5 region of the 18S rRNA gene. The species detected via visual methods are compared with those found via metabarcoding. Plankton communities are sampled at each of the same three locations with an Apstein net, once in June 2020 and once in August 2020. Morphological examination, which is likely to result in few identifications at a low taxonomic level, is not performed on the plankton samples; instead, they are directly metabarcoded with both the V4-V5 18S rRNA gene region and the COI gene region. For the plankton communities, the outputs of the two barcode gene approaches are compared. DNA sequencing is performed with a nanopore instrument (MinION of Oxford Nanopore Technologies, UK), allowing assessment of this recent advancement in sequencing technology for NIS detection.

Materials and Methods

I. Settlement plate communities

I.1: Sampling

Settlement plates were deployed to sample benthic communities at three locations in the harbour of Ostend on the 16th of June 2020: Vuurtorendok (51.237634°N, 2.931726°E), Ponton Overzet (51.234216°N, 2.927157°E), and Marina Mercator (51.228220°N, 2.923981°E) (Figure 2). Settlement plates were made of sanded grey PVC (15 * 15 cm) and suspended in the water column by means of a nylon rope that passed through a hole in the middle of each plate and was tied on one end to a pier and on the other end to a brick. In this manner, one or two settlement plates were deployed per location, so that the uppermost plate was suspended at 1 m depth and the lower plate was suspended at 7 m in accordance with the OSPAR-HELCOM protocol for NIS monitoring (OSPAR, 2019). The limited depth at Marina Mercator allowed the deployment of only one settlement plate (at 1 m depth), while two plates were deployed at each of the other locations. After two months (on 17 – 20 August 2020), all settlement plates were recovered and brought to the laboratory for analysis. First, a preliminary screening of live organisms visible on each side of each settlement plate was carried out to morphologically identify species (especially tunicates) whose morphology may be largely destroyed after preservation. Next, the benthic community attached to each side of each settlement plate was scraped off with a steel blade and preserved in DESS preservative.

DESS, an aqueous solution of 20 % dimethyl sulfoxide, 0.25 M EDTA, and saturated NaCl at pH 8.0 (Sharpe et al., 2020), is the recommended preservative for bulk marine communities due to its excellent ability to preserve DNA (van der Loos & Nijland, 2020). In addition, its shipping poses fewer risks than the highly flammable ethanol, which has traditionally been used for DNA preservation (van der Loos & Nijland, 2020). Note that in a number of samples, tunicates were extremely abundant, and some of these individuals had to be excluded from the preserved samples to facilitate morphological examination.

The samples of the settlement plates will be referred to with the abbreviation of their location (PO for Ponton Overzet, MM for Marina Mercator, and VD for Vuurtorendok) followed by the position (*Up* for upper plate and *Lo* for lower plate) and side (*Up* for upper side and *Un* for underside), respectively, of the corresponding plate (see Table 1). Hence, PO-LoUp stands for the upper side of the lower plate from Ponton Overzet, with other samples labelled likewise.

Figure 4: The three sampling locations in the harbour of Ostend. Credit: Google Maps, 2021.

Table 1: Naming convention for the settlement plate samples.

Location	Sample	Code
Ponton Overzet	Upper plate, upper side	PO-UpUp
	Upper plate, underside	PO-UpUn
	Lower plate, upper side	PO-LoUp
	Lower plate, underside	PO-LoUn
Vuurtorendok	Upper plate, upper side	VD-UpUp
	Upper plate, underside	VD-UpUn
	Lower plate, upper side	VD-LoUp
	Lower plate, underside	VD-LoUn
Marina Mercator	Upper plate, upper side	MM-UpUp
	Upper plate, underside	MM-UpUn

I.2: Morphological species identification

Following preservation, each sample (corresponding to one side of a settlement plate) was investigated under a stereomicroscope to identify as many taxa as possible to as low a taxonomic level as possible. Morphological identification of the preserved organisms was based on Hayward & Ryland (2017). Afterwards, subsamples for metabarcoding were taken from each sample. At least one specimen of each organism identified to species level was preserved as a voucher specimen.

I.3: DNA extraction

The DNA extraction protocol was largely based on that of the Global ARMS (Autonomous Reef Monitoring Structures) Project of the Smithsonian Institution (https://www.oceanarms.org). Each sample was homogenized in a blender and then poured over a 40 µm mesh net to retain the solid fraction. The solid fraction was subsequently mixed in a petri dish with a spatula before three subsamples of 0.25 g each (or as many subsamples as were possible if the solid fraction had a total mass below 0.75 g) were taken for DNA extraction. The blender, net, spatula, and petri dish were cleaned with a bleach solution after every use to avoid cross-contamination among samples. DNA was extracted from one subsample, with the exception of the two most speciose samples (PO-LoUp and VD-UpUn), of which all three subsamples were extracted. The latter step was taken to ensure that the least abundant species also had a fair chance of being detected via the metabarcoding. DNA extraction was carried out with the DNeasy® PowerSoil® Pro kit of Qiagen, with the added step of incubating each subsample in 50 µL of a 10 mg mL^{-1} proteinase K solution overnight at 56 °C after the addition of solution CD1 from the kit.

Once the DNA was extracted, its quality was checked with agarose gel electrophoresis and the NanoDrop® ND-1000 UV-Vis spectrophotometer (Thermo Scientific). The results indicated that contamination was present in the extracted genomic DNA (Table S1), with a 260/280 absorbance ratio of 1.42 – 2.63 and a 260/230 absorbance ratio of 0.05 – 0.79.

II. Zooplankton communities

II.1: Sampling

In June of 2020, zooplankton samples were collected at each of the three locations mentioned previously using a vertical free-fall plankton drop net (Apstein net) with a mesh size of 10 µm. The net was dipped three times at each location, and the plankton retrieved at each dip was first poured over a 300 µm mesh and then over a 100 µm mesh, yielding two size categories per replicate per location. Plankton samples were immediately preserved in 30-40 ml of DESS upon collection. This sampling design was repeated at the same three locations in August of 2020 while the settlement plates were being retrieved. In total, 36 zooplankton samples were taken.

II.2: DNA extraction

Approximately 10 mL (one quarter) of each sample was centrifuged for 1 min at 1000 rcf to precipitate the plankton organisms. After decantation of the supernatant, the pellet was resuspended in 10-11 mL of distilled water until all visible traces of salt crystals had vanished from the fluid. Each sample was again centrifuged and the supernatant discarded; the remaining pellet (containing the zooplankton) was then subjected to CTAB DNA extraction following a modified protocol of Cullings (1992). First, 300 µL of CTAB (cetrimonium bromide) buffer was added to the pellet. This buffer was prepared as in Cullings (1992), except that polyvinylpyrrolidone was not added. Next, the pellet was transferred to a vial containing 350 mg of 0.1 mm glass beads and 350 mg of 0.5 mm glass beads. All vials were shaken for 3 x 30 s on a Minilys homogenizer (Bertin Technologies, Montigny-le-Bretonneux, France) with 1 min of cold block incubation between consecutive runs to avoid overheating the samples. Next, 20 µL of 10 mg/mL proteinase K solution was added to each vial and all vials were incubated with rotation for one hour at 55°C. Afterwards, the lysate was transferred out of each vial into a new microcentrifuge tube and mixed with 300 µL of chloroform. The tubes were vortexed and centrifuged for 15 min at 14,000 rcf. The aqueous phase of each sample was transferred to a new tube and again mixed with 300 µL of chloroform, vortexed, and centrifuged for 15 min at 14,000 rcf.

The resulting aqueous phase of each sample was again transferred to a new microcentrifuge tube, this time together with 60 µL 3M sodium acetate and 350 µL isopropanol. After vortexing and centrifugation (1 min at 14000 rcf), 300 µL of 70 % ethanol was added to each sample. Following one minute of incubation at room temperature, the microcentrifuge tubes were centrifuged for one minute at 14,000 rcf, drained, and allowed to dry at room temperature for 15 min. The genomic DNA, which had formed a pellet adhering to the lower end of each tube, was resuspended in 30 µL of Tris-EDTA buffer for storage.

The quantity of extracted DNA from the plankton samples was measured using the Invitrogen Qubit 3 Fluorometer (ThermoFisher Scientific). In addition, a preliminary PCR was performed on a subset of samples (as well as on a negative control) using the F-566 and R-1200 primer pair, which targets the V4-V5 region of the 18S rRNA gene (Hadziavdic et al., 2014). This preliminary PCR was performed to verify that the genomic DNA was of sufficient quality to be PCR-amplified, even when the fluorometer measurements indicated low DNA concentration.

III. PCR amplification, library preparation and sequencing

All samples from the zooplankton and the settlement plates were metabarcoded twice, once with the F-566/R-1200 primer pair (which targets the V4-V5 18S rRNA gene region; Hadziavdic et al., 2014) and another time with the dgLCO1492/dgHCO2198 pair (which are degenerate primers targeting the COI gene; Meyer et al., 2005). The jgLCO1492/jgHCO2198 (Geller et al., 2013) and Lobo1F/Lobo1R primers (Lobo et al., 2013) were originally tried for the COI regions, but the amplification success of both primer pairs was very poor. For each set of primers, two consecutive PCRs were carried out on the samples before attachment of sequencing adapters. In the first reaction, the region of interest was amplified with the aforementioned primer pairs. Primers had a 5'-tail to which primers of the Oxford Nanopore Technologies PCR Barcoding Kit (SQK-PBK004) would bind in the second PCR reaction in order to label individual samples with unique barcodes for the purpose of multiplexing. Each reaction was followed by a clean-up using paramagnetic beads (CleanPCR from GC Biotech, Waddinkxveen, the Netherlands) and then a quality check using the Qubit 3 fluorometer (ThermoFisher Scientific). For the first PCR, 25 µL reactions were carried out using 12.5 µL Phire Hot Start II polymerase (2X) (ThermoFisher Scientific), 9.5 µL distilled water, 1 µL of forward primer, 1 µL of reverse primer, and 1 µL (approximately 0.2 ng) of purified DNA from the sample. Three PCR reactions were prepared per sample in order to account for possible amplification biases; the three reaction products were subsequently pooled before the first clean-up. The following settings were used for the first PCR: 60 s initial denaturation at 98°C; 30 cycles of denaturation (5 s at 98°C), annealing (15 s at 60°C), and extension (15 s at 72°C); and a final extension of 7 min at 72°C. The second PCR was used to attach multiplexing barcodes to the reads; the barcodes used were those of the Oxford Nanopore Technologies

PCR Barcoding Kit (SQK-PBK004). For the second PCR, 50 µL reactions (each corresponding to one sample) were carried out using 25 µL Phire Hot Start II polymerase (2X) (ThermoFisher Scientific), 1 µL of barcode primer, and exactly 10 ng previously amplified DNA. The reactions were topped off with distilled water. The following settings were used for the second PCR: 60 s initial denaturation at 98°C; 14 cycles of denaturation (5 s at 98°C), annealing (15 s at 56°C), and extension (15 s at 72°C); and a final extension of 7 min at 72°C. After the final clean-up, all barcoded libraries were pooled to a total of 25 to 50 ng (in 10 µL of a solution consisting of 10 mM Tris-HCl and 50 mM NaCl at pH 8.0) before ligation of rapid adapters following the protocol of Oxford Nanopore Technologies. Because no more than 12 unique multiplexing barcodes from the SQK-PBK004 could be used at once, the samples analysed in this study had to be split into two runs. In the first run, all samples from the settlement plates plus an extra replicate from Ponton Overzet (one of the two most speciose samples) were sequenced alongside a negative control that had likewise been subjected to the same PCR amplification protocol. In the second run, all plankton samples (both 18S and COI) were sequenced. Note that due to constraints in time and resources, not all extracted samples could be sequenced.

The barcoded libraries were sequenced on the MinION sequencer (Oxford Nanopore Technologies, UK, 2021). In the first run, 680,020 reads were generated on a FLO-MIN111 flow cell. In the second run, 7.53 million reads were generated on a FLO-MIN106 flow cell.

IV. Data analysis and visualization

In both cases, the MinKnow program v.21.02.5 (Oxford Nanopore Technologies, United Kingdom, 2021) performed automatic basecalling of the reads as they were produced (using the program's Fast Basecalling model implemented in Guppy v.4.3.4 (Oxford Nanopore Technologies, United Kingdom, 2021)). Nanopore barcoding sequences and 5'-tail sequences of primers were trimmed using MinKnow.

The FASTQ files resulting from the sequencing were analysed via the Decona version 0.1.2 pipeline (https://github.com/Saskia-Oosterbroek/decona), which was developed specifically for the long reads typical of Nanopore sequencing. For both runs, sequences were selected with a minimum quality threshold of 8 and a length range of 500 – 1000 bp, with a minimum percentage match to the database sequences of 80 %. Clustering was performed using the more accurate of the two algorithms available in Decona, and polishing was performed with Medaka. The reads were aligned to the SILVA 138.1 SSU Ref_NR99 database (for the 18S rRNA reads of the settlement plates and plankton) and the MetaZooGene database (MZGdb; Bucklin et al., 2021) (for the COI reads of the plankton).

NMDS (nonmetric multidimensional scaling) was used to visualize similarities across samples analysed with a given method. These data visualizations were performed using the

metaMDS() function from the 'vegan' package (version 2.5-7; Oksanen et al., 2020) in R version 3.6.1. Dissimilarity matrices were calculated for each dataset using the vegdist() function of this package, and these matrices can be found in the supplementary material (Figures S2-S5). For the settlement plates, NMDS was performed using the Raup-Crick index (Raup & Crick, 1979) as a distance measure for the morphological data and the Bray-Curtis dissimilarity measure (Bray & Curtis, 1957)) for the sequencing data. For the plankton communities, the decostand() function from *vegan* was first employed to standardize the read counts in the plankton sequencing data in order to avoid having the community variation dominated by a few, highly abundant species. Afterwards, NMDS was conducted using Euclidean distances of the standardized data. The NMDS plots and bar charts of sequence abundances were created using 'ggplot2' (version 3.3.3; Wickham, 2016) and 'ggrepel' (version 0.9.1; Slowikowski, 2021).

Results

I. Morphological identification of settlement plate communities

Morphological examination of the communities on the settlement plates resulted in the discovery of at least 36 taxa (Table 2), of which 16 could be identified to species level. The estimate of 36 taxa is based on the assumption that taxa above the species level included only one species, unless various individuals were found belonging to different species within this taxon. In this case, the higher taxon was assumed to contain only those species. For example, some pycnogonids (sea spiders) were assigned to the species *Anoplodactylus virescens*, while others were identified to genus or order (Table 2). In this case, it is considered that one species of pycnogonid is present in the samples, even though multiple species may still have been present. In this way, the estimate of 36 taxa is considered conservative.

Sixteen entries (taxa or at least morphological groups) are unique to one sampling location (Table 2). These are *Maera grossimana*, Natantia, Varunidae, Tanaidae, *Chaetomorpha linum*, both *Pedicellina* species, and rhodophytes for Ponton Overzet; *Idotea* sp., the unidentified barnacles and pycnogonids, *Clavelina lepadiformis*, Stauromedusae, and the unidentified bivalves for Vuurtorendok; and *Anoplodactylus* sp. and *Obelia longissima* for Marina Mercator. The NMDS plot for the samples does indeed suggest location-specific communities structured along the first axis (Figure 5). The identity of individual settlement plates (upper or lower plate) and side of the plate seemed to have little effect on the sample clustering (Figure 5). The stress value of the NMDS was 0.7.

Austrominius modestus (New Zealand barnacle) and the tunicate *Ciona intestinalis* were the only non-native taxa identified to species level. *A. modestus* is listed on the OSPAR-HELCOM list of non-indigenous species of concern for the North Sea (http://jointbwmexemptions.org). *Diplosoma listerianum, Amphibalanus improvisus,* and *Monocorophium sextonae* are cryptogenic species (Hayward & Ryland, 2017; Verleye et al., 2020) and not included in the OSPAR-HELCOM list. Specimens of the crab family Varunidae could not be identified to genus or species level. This family is represented in northwest Europe exclusively by introduced species (Hayward & Ryland, 2017), three of which (*Eriocheir sinensis*, *Hemigrapsus sanguineus*, and *H. takanoi*) are on the OSPAR-HELCOM list. The genera *Botrylloides, Molgula,* and *Caprella* and the families Serpulidae and Tanaidae include both native and non-native species (Hayward & Ryland, 2017; Verleye et al., 2020), with some of the latter occurring on the OSPAR-HELCOM watch list. Within the Natantia, *Palaemon macrodactylus* (Oriental prawn) is a NIS documented for the Belgian part of the North Sea (Verleye et al., 2020). Bryozoans, bivalves, rhodophytes, and chlorophytes also include various North Sea representatives on this watch list, but except for *Chaetomorpha linum*, a native, could not be accurately identified any further than these groupings. Those copepods, barnacles, decapod larvae, and ascidians that could not be accurately classified are likewise of dubious origin.

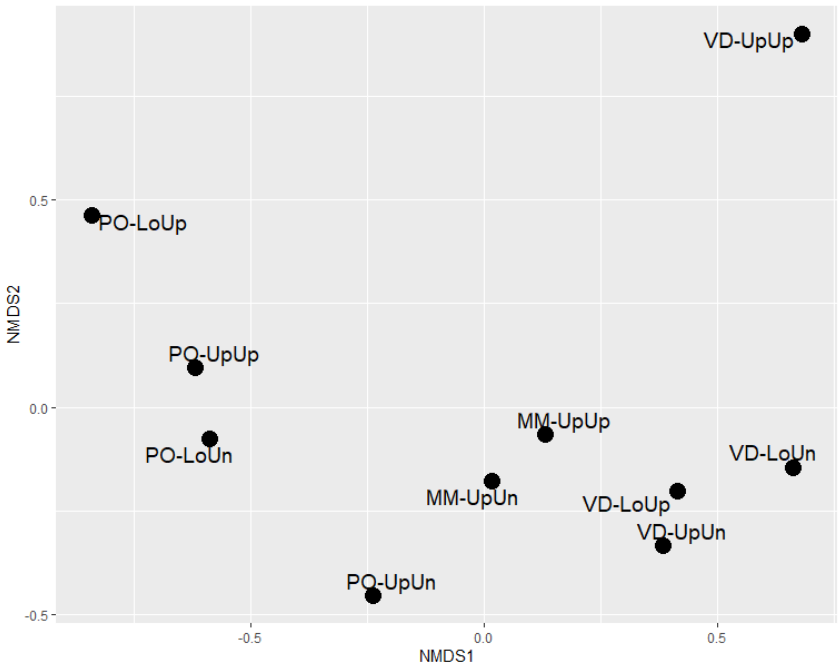

Figure 5: NMDS plot for the morphologically identified settlement plate samples. For the abbreviations, see Table 1.

Table 2: Taxa found via morphological examination of settlement plate communities. Species known to be introduced are marked with an asterisk. Genera and families containing both native species and confirmed introduced species are marked with an asterisk in parentheses. Sample names follow the convention in the text and in Table 1.

Phylum	Class	Order	Name	PO-UpUp	PO-UpUn	PO-LoUp	PO-LoUn	VD-UpUp	VD-UpUn	VD-LoUp	VD-LoUn	MM-UpUp	MM-UpUn
Annelida	Polychaeta	Phyllodocida	Phyllodocidae sp.	x	x	x	x		x	x		x	x
			Syllidae sp.						x	x	x	x	
		Sabellida	Serpulidae sp. (*)	x	x	x	x		x	x	x	x	x
Arthropoda	Hexanauplia	Cyclopoida	*Halicyclops* sp.		x			x		x			
		NA	(unidentified copepods)			x		x			x	x	
	Malacostraca	Amphipoda	*Abludomelita obtusata*		x		x		x				
			Caprella sp. (*)		x	x			x	x	x	x	x
			Crassicorophium bonellii				x		x		x	x	x
			Maera grossimana			x							
			Melitidae sp. (*)			x		x			x		x
			Microdeutopus gryllotalpa					x	x	x	x	x	x
			Monocorophium sextonae	x	x	x	x	x	x	x		x	x
		Decapoda	Natantia sp.				x						
			Varunidae sp. *			x							
			(unidentified decapod larvae)		x	x							x
		Isopoda	*Idotea* sp.					x	x	x	x		
		Tanaidacea	Tanaidae sp. (*)			x							
	Thecostraca	Balanomorpha	*Amphibalanus improvisus*			x	x					x	x
			Austrominius modestus *	x		x	x		x		x		
			(unidentified barnacle)					x					
	Pycnogonida	Pantopoda	*Anoplodactylus* sp.										x
			Anoplodactylus virescens						x			x	
			(unidentified pycnogonid)							x			
Bryozoa	NA	NA	(branched bryozoan)	x	x	x	x		x	x	x	x	x

Table 2, continued

Phylum	Class	Order	Name	PO-UpUp	PO-UpUn	PO-LoUp	PO-LoUn	VD-UpUp	VD-UpUn	VD-LoUp	VD-LoUn	MM-UpUp	MM-UpUn
			(crust-like bryozoan)			x	x				x		
Chlorophyta	Ulvophyceae	Cladophorales	*Chaetomorpha linum*			x							
	NA	NA	(branched, filamentous green algae)	x		x							
	NA	NA	(foliose and/or tubular green algae)	x	x		x					x	x
Chordata	Ascidiacea	Aplousobranchia	*Clavelina lepadiformis*						x	x	x		
			Diplosoma listerianum						x		x		x
		Phlebobranchia	*Ascidiella* sp.		x				x	x	x	x	x
			*Ciona intestinalis**		x				x		x	x	x
		Stolidobranchia	*Botrylloides* sp. (*)		x				x		x		
			Botryllus schlosseri	x		x	x		x	x	x	x	x
			Molgula sp. (*)							x		x	x
		NA	(unidentified solitary tunicate)	x		x	x						
Cnidaria	Hydrozoa	Leptothecata	*Obelia longissimi*									x	
	Staurozoa	Stauromedusae	Stauromedusae sp.								x		
Entoprocta	NA	NA	*Pedicellina cernua*		x		x						
			Pedicellina hispida			x							
Mollusca	Bivalvia	NA	(unidentified bivalve)					x					
Nematoda	NA	NA	(unidentified nematodes)			x						x	
Rhodophyta	NA	NA	(branched, filamentous red algae)	x	x	x							

II. Metabarcoding of settlement plate and plankton communities

II.1: Summary of the MinION sequences

The first run of the sequences continued for 2 d 21 h 21 min, generating a total of 680,020 reads. The second run took 12 h 2 min, generating 7.53 million reads. In the first run, 406.95 million bases (Mb) passed the quality control, while 117.09 Mb failed. In the second run, 4.22 Gb and 329.42 Mb passed and failed, respectively. The estimated total number of bases for the runs were 576.44 Mb and 4.7 Gb for the first and second runs, respectively. Figure S6 shows that the vast majority of reads had a length of around 750 bp, corresponding to the 18S rRNA gene fragment (including primers, multiplexing barcodes, and the 5' tails). In Figure S7, two major peaks of around 750 bp and 850 bp corresponding to 18S and COI gene regions, respectively, can be noted. The proportion of reads and bases of each barcode that passed the sequencing run can be found in Figures S8 and S9. Typically, approximately three-fourths or more of the reads and bases passed the runs, with the proportion passing being higher for the second run. The quality score of the reads generally stayed within the median target for the duration of both runs, staying around 8 in the first run (Figure S10) and slightly declining from 10 in the second run (Figure S11).

II.2: 18S rRNA metabarcoding of settlement plate communities

Metabarcoding of the 18S rRNA gene region recovered 48 taxa from the settlement plates, with 32 taxa resolved to species level and 10 resolved to genus level (Table 3). Sampling locations showed many location-specific taxa, with only 10 of the taxa in Table 3 being found in more than one of the three locations. Again, the three locations could be differentiated on the first axis of the NMDS plot (stress value 0.7; Figure 6), although here the Marina Mercator samples tended to cluster together with the Vuurtorendok samples. This may be due to the many reads of ascidians found at both locations (Table 3). The percentage alignment of the sequences to the SILVA database ranged from 81.5 % to 100 %. The negative control also contained sequences, but these were almost all from organisms never detected (either morphologically or via metabarcoding) in any of the settlement plate samples (such as humans and various crop plants). The only exception was *Austrobalanus imperator*, which was present in the negative control as well. All sequences unique to the negative control are therefore excluded from table 3. In addition, the terrestrial desert alga *Acutodesmus deserticola* was represented by 20 reads in sample PO-UpUp and probably represents either a contamination or a mismatch; hence, it was removed from the inventory as well.

Among the taxa detected, only *Ficopomatus enigmaticus* is on the OSPAR-HELCOM watch list. Verleye et al. (2020) list *Amphibalanus amphitrite*, *Ficopomatus enigmaticus*, and

Botrylloides violaceus as NIS in the Belgian part of the North Sea. 15 taxa are recorded from the Belgian part of the North Sea for the first time (*Paramphiascella fulvofasciata*, *Ianiropsis epilittoralis*, *Sclerochilus oshoroensis*, *Austrobalanus imperator*, *Striatobalanus amaryllis*, *Ascidia ceratodes*, the two *Molgula* species, *Styela plicata*, *Symplegma viride, Lankesteria halocynthiae*, *Thalassiosira allenii*, *Haliphthoros* sp., *Kaitalugia* sp., and *Diapse ptilota*). The Bicosoecida reads were matched to those of freshwater species in the sequence database, while the metadata of the Nucletmycea reads listed them as belonging to a fungal metagenome.

Interestingly, the two samples of the upper side of the lower plate from Ponton Overzet (PO-LoUp) shared only two taxa (*A. imperator* and Chromadorida), with seven taxa unique to the first replicate and only one taxon unique to the second (see table 3). Of the seven taxa unique to the first replicate, six were represented by less than 100 reads, representing a possible coverage issue (see table 6). The majority of reads from the settlement plates were assigned to either arthropods or chordates (in this case exclusively ascidians), although diatoms (corresponding to Ochrophyta) were relatively well-represented on the upper side of the upper plate from Ponton Overzet (Figure 7).

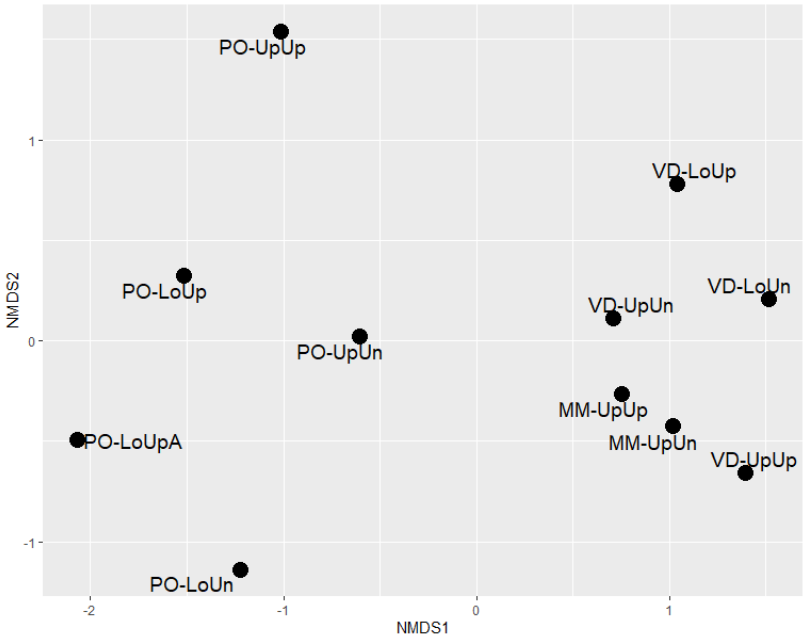

Figure 6: NMDS plot of the settlement plate samples based on 18S rRNA metabarcoding. For the abbreviations, see Table 1.

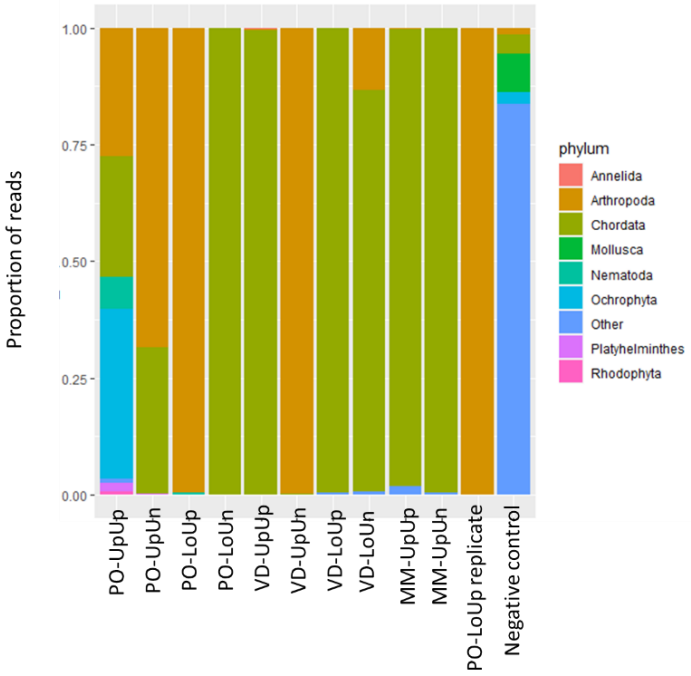

Figure 7. Proportion of reads assigned to different phyla from the settlement plates. For the abbreviations, see Table 1.

Table 3: Taxa detected on the settlement plates via 18S rRNA metabarcoding. Species marked with an asterisk are known to be introduced. Taxa marked with an asterisk in parentheses have not been recorded from the Belgian part of the North Sea yet. Sample names follow the convention in the text. For site abbreviations, see Table 1.

Phylum	Class	Order	Name	Max. alignment (%)	Total number of reads per taxon										
Location Sample					PO-UpUp	PO-UpUn	PO-LoUp	PO-LoUp	PO-LoUn	VD-UpUp	VD-UpUn	VD-LoUp	VD-LoUn	MM-UpUp	MM-UpUn
Annelida	Polychaeta	Sabellida	*Ficopomatus enigmaticus* *	99.55								148			
Arthropoda	Hexanauplia	Calanoida	*Centropages hamatus*	100	43										
		Cyclopoida	*Lichomolgus canui*	100		23									
		Harpacticoida	*Ameira scotti*	98.607			46								
			Argestidae sp.	92			101								
			Argestigens sp.	96.279					31						
			Bradya sp.	94.072			26								
			Harpacticus sp.	98.762		47									
			Paramphiascella fulvofasciata (*)	99.065						21					
			Tisbe sp.	95.956						101					
	Malacostraca	Isopoda	*Ianiropsis epilittoralis* (*)	98.378								20			
	Ostracoda	Podocopida	*Sclerochilus oshoroensis* (*)	97.962						103				21	66
			Xiphichilus sp.	98.583				21							
	Thecostraca	Balanomorpha	*Amphibalanus amphitrite* *	99.705			22			4489					
			Austrobalanus imperator (*)	99.558	6998	26798	35482	29177	44857						
			Semibalanus balanoides	99.41			84								
			Striatobalanus amaryllis (*)	99.134	27	27									
Bryozoa	Gymnolaemata	Ctenostomatida	*Amathia* sp.	100	26										
Bigyra	Bicoecea	NA	Bicosoecida sp. (*)	83.168	30	27									
Chlorophyta	Ulvophyceae	Ulvales	*Ulva rigida*	99.681	138										
Chordata	Ascidiacea	Phlebobranchia	*Ascidia ceratodes* (*)	95.32											22

Phylum	Class	Order	Name	Max. alignment (%)	Total number of reads per taxon										
Location Sample					PO-UpUp	PO-UpUn	PO-LoUp	PO-LoUp	PO-LoUn	VD-UpUp	VD-UpUn	VD-LoUp	VD-LoUn	MM-UpUp	MM-UpUn
			Ascidiella sp.	100		2178			47	30505	544	36117	4142	15266	31863
			Ciona intestinalis *	100							23163		29130	1601	
		Stolidobranchia	*Botrylloides violaceus* *	100		372				22	111		2450		
			Botryllus schlosseri	99.064	52		21				20	230		30	
			Molgula provisionalis (*)	100	704	9672					26				
			Molgula retortiformis (*)	99.842		48			32					7262	553
			Styela plicata (*)	99.373	5888										
			Symplegma viride (*)	91.615	26										
NA	NA	NA	Nucletmycea sp.	82.012						240	23			42	298
Nematoda	Chromadorea	Chromadorida	Chromadorida sp.	98.899	135		68	36							
			Chromadorina sp.	100	446										
			Chromadorita leuckarti	98.259						20					
			Chromadorita tentabundum	91.875	64										
			Punctodora ratzeburgensis	93.049	316										
	Enoplea	Enoplida	*Viscosia* sp.	99.842	678		80								
	NA	NA	Nematoda sp.	91.85	159										
Myzozoa	Conoidasida	Eugregarinorida	*Lankesteria halocynthiae* (*)	91.613		48					79			59	284
Ochrophyta	Bacillariophyceae	Coscinodiscales	*Actinoptychus splendens*	98.273	30										
		Chaetocerotanae	*Chaetoceros socialis*	99.53	149										
		Melosirales	*Melosira dubia*	96.378	9090										
		Thalassiosirales	*Thalassiosira allenii* (*)	98.442	57										
			Thalassiosira angulata	97.174	37										
Oomycota	Peronosporea	Saprolegniales	*Haliphthoros* sp. (*)	91.859	20										
Platyhelminthes	Turbellaria	Rhabdocoela	*Kaitalugia* sp. (*)	98.532	462	26									
		Prolecithophora	*Pseudostomum quadrioculatum*	82.753		32									
Rhodophyta	Florideophyceae	Ceramiales	*Ceramium rubrum*	98.728	143										
			Diapse ptilota (*)	96.513	46										

Comparing the results from the morphological examination with those of 18S rRNA metabarcoding is challenging because of the varying levels of taxonomic resolution achieved across the samples. However, when comparing organisms identified to species level, it becomes clear that the two methods are highly divergent, with only two species (*Ciona intestinalis* and *Botryllus schlosseri*) being recorded by both methods (Figure 8).

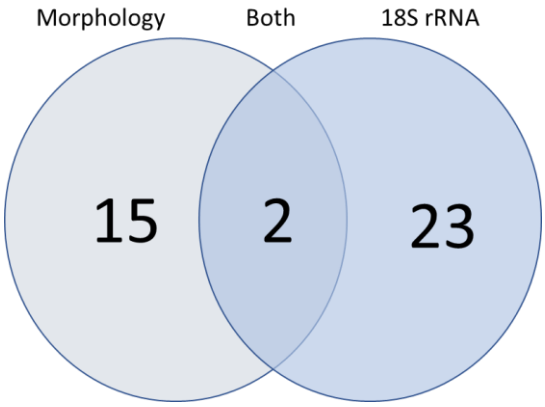

Figure 8: Comparison of the number of species detected from the settlement plates using either identification method (only species with an alignment greater than 97% are considered for the 18S rRNA data).

II.3 18S rRNA metabarcoding of plankton communities

In total, 18S rRNA metabarcoding of all plankton samples from both months yielded 37 taxa, 29 of which were assigned to a particular species or putative species (see Tables 4 and 5). Of the remaining eight taxa, seven were resolved to genus level (Tables 4 and 5). Only two species (*Hemiaulus sinensis* and *Polydora haswelli*) were sampled at both time points. Of the June 2020 samples, four taxa (*P. haswelli*, *Acartia clausii*, *Noctiluca scintillans*, and *Chaetoceros* sp.) were found at more than one sampling location (Table 4), while of the August 2020 samples, six taxa (*Bellerochea malleus* and all copepods except *Mesochra* sp. and *Thompsonula hyaenae*) were found at more than one location (Table 5). Percentage alignment to the SILVA database ranged from 88.2 % to 100 %.

The NMDS plot (stress value 0.06 with two convergent solutions) recognized a cluster of seven locations in addition to five more divergent locations that did not cluster as well (Figure 9). The prevalence of diatoms (Figure 10), which were not targeted by the sampling strategy (due to the mesh sizes used) but probably included due to plankton net clogging or colony formation, may explain the positioning of the two samples on the right-most part of axis 1 of Figure 9. Conversely, the large proportion of tunicate reads from Marina Mercator (Figure 10) seems to have placed this location to the left extremity of the first axis (Figure 9).

Acartia tonsa, *Magallana gigas*, *Ficopomatus enigmaticus*, and *Pseudodiaptomus marinus* are all recognized NIS in the Belgian part of the North Sea according to Verleye et al. (2020); *Ciona intestinalis* and *Pseudopolydora paucibranchiata* are also considered alien (Horton et al., 2021). *A. tonsa*, *C. gigas*, and *F. enigmaticus* are included on the OSPAR-HELCOM watch list. *P. haswelli*, *Molgula retortiformis*, *Pseudopolydora paucibranchiata*, *Hemiaulus sinensis*, *Minidiscus* sp., *Skeletonema grethae*, *Chaetoceros lorenzianus*, *Thalassiosira oceanica*, both *Oithona* species, and *Tachidius triangularis* have not yet been described from the Belgian part of the North Sea. The proportion of reads per phylum from the June samples varies greatly per location; tunicates make up a substantial portion only in Marina Mercator, *Noctiluca*

scintillans (the only representative of Myzozoa found) forms a sizeable fraction of the reads only at Ponton Overzet, and the reads from Vuurtorendok are dominated by diatoms (phylum Ochrophyta) (figure 10). By August 2020, however, all samples came to be dominated by reads from Arthropoda, which was solely represented by copepods (figure 11).

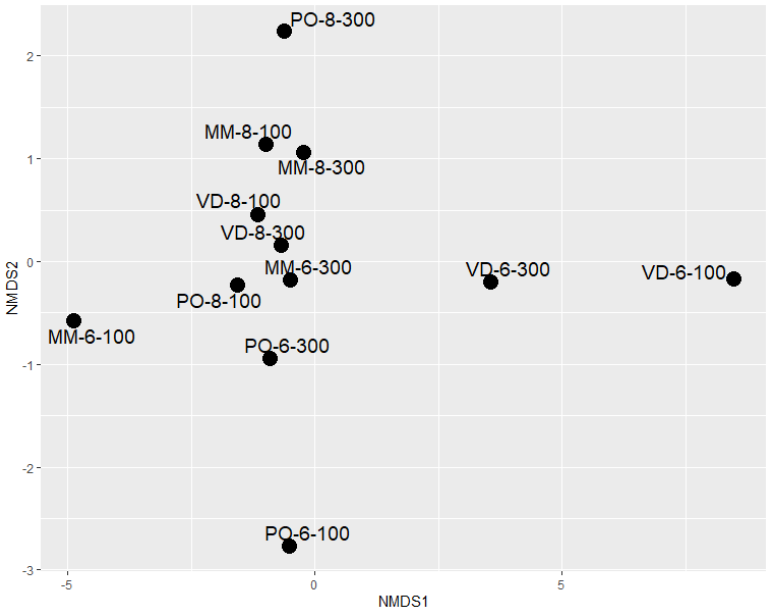

Figure 9: NMDS plot of the plankton samples based on 18S rRNA metabarcoding. Sample names are given as abbreviation of location, followed by month of sampling and then mesh size.

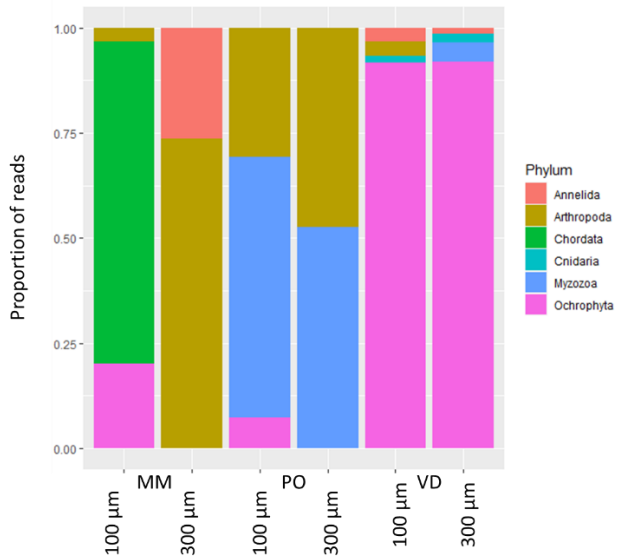

Figure 10: Proportion of 18S rRNA reads per phylum from the June 2020 plankton samples, with location and mesh size indicated.

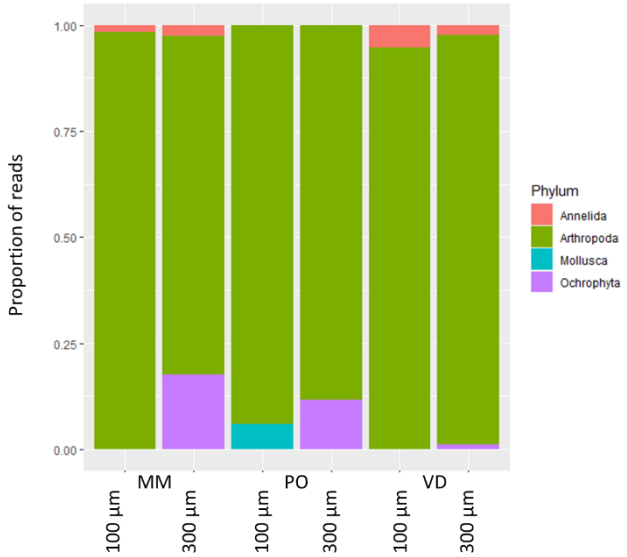

Figure 11: Proportion of 18S rRNA reads per phylum from the August 2020 plankton samples, with location and mesh size indicated.

II.4: COI metabarcoding of plankton communities

Twenty-three planktonic taxa were detected via COI metabarcoding (Tables 6 and 7), all of which were assigned to a particular species. Percentage alignment to the database ranged from 80.1 % to 100 %. The three species *Athorybia rosacea*, *Boccardia pugettensis* and *Obelia dichotoma* were shared between the two months, while all other species were detected only in samples from one of the two months (Tables 6 and 7). *O. dichotoma* was the only species found in more than one sampling location in June 2020 (Table 6), while in August 2020, three species (*Boccardia pugettensis*, *Kirchenpaueria pinnata*, and *Magallana gigas*) were found in more than one location (table 7). Despite the low number of taxa shared across locations and dates, the NMDS plot (which had a stress value of 0.13 and two convergent solutions; Figure 12) weakly clustered the Vuurtorendok samples in the upper left, the Marina Mercator samples in the lower right, and the Ponton Overzet samples in the lower left, indicating a certain level of influence of location on community structure.

Only *Magallana gigas* is an acknowledged NIS in Belgian marine areas (Verleye et al., 2020), and it is also the only species detected by COI that is on the OSPAR-HELCOM watch list. However, the following species have not yet been recorded for the Belgian part of the North Sea: *Barrukia cristata*, *B. pugettensis*, *Platorchestia pachypus*, *Metapenaeus ensis*, *Portunus sanguinolentus*, *Stenasellus racovitzai*, *Ototyphlonemertes santacruzensis*, *Oithona davisae*, and *Ciliopagurus strigatus*. The proportion of reads per phylum appeared very different compared to those of the 18S rRNA reads. In the COI reads, arthropods dominated the June 2020 sequences (Figure 13) while annelids and cnidarians made up the bulk of the August 2020 sequences (Figure 14). One notable exception to this pattern is the sole June sample from Marina Mercator (Figure 13) (of which the 300 µm fraction yielded no assigned sequences). Here, rotifers, which remain undetected elsewhere in the plankton samples, make up about 67% of the reads. No ochrophytes or other metazoan taxa were detected via COI metabarcoding. In agreement with the 18S rRNA data, however, tunicate sequences were detected exclusively in the June samples from Marine Mercator (Figure 13).

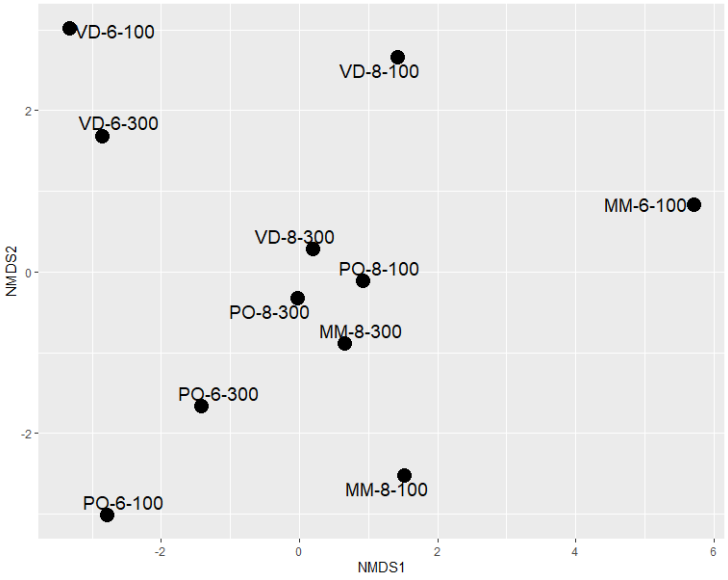

Figure 12: NMDS plot of the plankton samples based on COI metabarcoding. Sample names are given as abbreviation of location, followed by month of sampling and then mesh size.

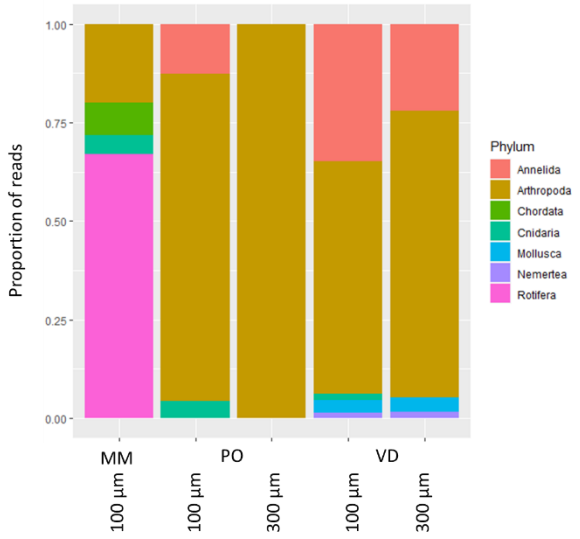

Figure 13: Proportion of COI reads per phylum from the June 2020 plankton samples, with location and mesh size indicated.

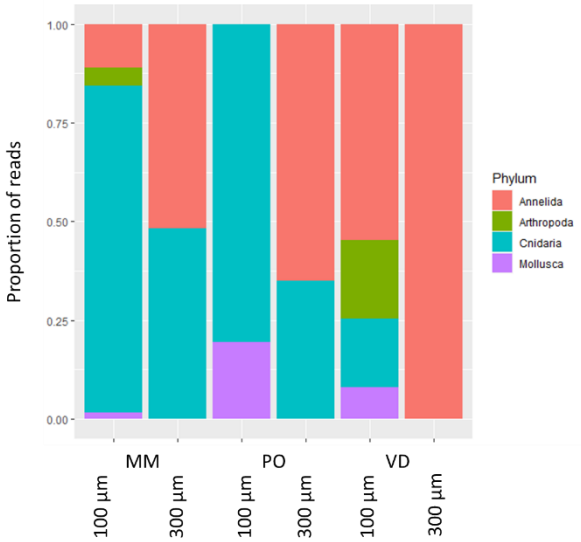

Figure 14: Proportion of COI reads per phylum from the August 2020 plankton samples, with location and mesh size indicated.

Table 4. Taxa detected in the June 2020 plankton samples via 18S metabarcoding. Species marked with an asterisk are known to be introduced. Species marked with an asterisk in parentheses have not been recorded from the Belgian part of the North Sea yet.

Phylum	**Class**	**Order**	**Name**	**Max. alignment (%)**	**Reads per taxon**					
Location					PO	PO	MM	MM	VD	VD
Mesh size					(100 µm)	(300 µm)	(100 µm)	(300 µm)	(100 µm)	(300 µm)
Annelida	Polychaeta	Spionida	*Polydora haswelli* (*)	99.068				150	247	117
Arthropoda	Hexanauplia	Calanoida	*Acartia clausii*	99.685	1891	1803		417		
			Calanoida sp.	99.37	111	102				
			Centropages hamatus	99.688	232	106				
		Cyclopoida	*Lichomolgus canui*	98.752					264	
			Pseudanthessius sp. (New Caledonia) (*)	98.746			142			
Chordata	Ascidiacea	Phlebobranchia	*Ascidiella* sp.	99.064			2578			
			Ciona intestinalis *	97.623			508			
		Stolidobranchia	*Molgula retortiformis* (*)	99.372			304			
Cnidaria	Hydrozoa	Leptothecata	*Obelia longissima*	99.696					126	148
Myzozoa	Dinophyceae	Noctilucales	*Noctiluca scintillans*	99.531	4496	2239				371
Ochrophyta	Bacillariophyceae	Chaetocerotanae	*Chaetoceros* cf. *lorenzianus* (*)	98.077					194	224
			Chaetoceros costatus	98.11						199
			Chaetoceros debilis	99.526					260	140
			Chaetoceros didymus	98.392					241	186
			Chaetoceros mitra	95.141					111	
			Chaetoceros radicans	99.527					3249	3438
			Chaetoceros sp.	95.584			889		418	
		Hemiaulales	*Hemiaulus sinensis* (*)	91.499					119	146
		Rhizosoleniales	*Guinardia striata*	99.523	526					

Phylum	Class	Order	Name	Max. alignment (%)	Reads per taxon					
Location Mesh size					PO (100 µm)	PO (300 µm)	MM (100 µm)	MM (300 µm)	VD (100 µm)	VD (300 µm)
		Thalassiosirales	*Lauderia borealis*	99.22					634	383
			Minidiscus sp. (*)	97.92					295	
			Skeletonema grethae (*)	99.369					434	103
			Thalassiosira anguste-lineata	98.74					675	396
			Thalassiosira oceanica (*)	98.122					228	1934
			Thalassiosira sp.	98.578					176	

Table 5: Taxa detected in the August 2020 plankton samples via 18S metabarcoding. Species marked with an asterisk are known to be introduced. Species marked with an asterisk in parentheses have not been recorded from the Belgian part of the North Sea yet.

Phylum	Class	Order	Name	Max. alignment (%)	Reads per taxon					
Location Mesh size					PO (100 µm)	PO (300 µm)	MM (100 µm)	MM (300 µm)	VD (100 µm)	VD (300 µm)
Annelida	Polychaeta	Sabellida	*Ficopomatus enigmaticus* *	99.407			142			
		Spionida	*Polydora haswelli* (*)	98.445				146		
			Pseudopolydora paucibranchiata *	99.69					465	229
Arthropoda	Hexanauplia	Calanoida	*Acartia tonsa* *	100	756			193		137
			Pseudodiaptomus marinus *	97.656		118		153		
		Cyclopoida	*Oithona davisae* (*)	98.752	342	332			6788	7485
			Oithona sp. 1 (New Caledonia) (*)	96.899		846	2659	1369	591	
		Harpacticoida	*Mesochra* sp.	92.581		176				

Phylum	Class	Order	Name	Max. alignment (%)	Reads per taxon					
Location Mesh size					PO (100 µm)	PO (300 µm)	MM (100 µm)	MM (300 µm)	VD (100 µm)	VD (300 µm)
			Tachidius triangularis (*)	95.054	1474	5516	6078	2612	707	1735
			Thompsonula hyaenae	88.208		164				
Mollusca	Bivalvia	Ostreida	*Magallana gigas* *	99.377	162					
Ochrophyta	Bacillariophyceae	Hemiaulales	*Bellerochea malleus*	99.686		597		823		118
			Hemiaulus sinensis (*)	95.469				139		

Table 6: Taxa detected in the June 2020 plankton samples via COI metabarcoding. Species marked with an asterisk are known to be introduced. Species marked with an asterisk in parentheses have not been recorded from the Belgian part of the North Sea yet.

Phylum	Class	Order	Name	Max. alignment (%)	Reads per taxon					
Location Mesh size					PO (100 µm)	PO (300 µm)	MM (100 µm)	MM (300 µm)	VD (100 µm)	VD (300 µm)
Annelida	Polychaeta	Phyllodocida	*Barrukia cristata* (*)	81.543					244	
		Spionida	*Boccardia pugettensis* (*)	84.615					2351	1745
		Terebellida	*Lanice conchilega*	84.992	403					
Arthropoda	Hexanauplia	Calanoida	*Pseudocalanus elongatus*	99.24	1209	4000				
	Malacostraca	Amphipoda	*Gammarus crinicornis*	98.321		267				
			Platorchestia pachypus (*)	80.628			480			
		Decapoda	*Lysmata seticaudata*	82.043	1440	622				
			Metapenaeus ensis (*)	83.684					267	
			Parasergestes vigilax	81.675			101			
			Portunus sanguinolentus (*)	86.466					110	128
		Isopoda	*Stenasellus racovitzai* (*)	84.066					4033	5658

Phylum	Class	Order	Name	Max. alignment (%)	Reads per taxon					
Location Mesh size					PO (100 µm)	PO (300 µm)	MM (100 µm)	MM (300 µm)	VD (100 µm)	VD (300 µm)
Cnidaria	Hydrozoa	Leptothecata	*Obelia dichotoma*	89.482			142		112	
		Siphonophorae	*Athorybia rosacea*	84.615	136					
Chordata	Ascidiacea	Phlebobranchia	*Ascidiella aspersa*	97.432			238			
Mollusca	Bivalvia	Mytilida	*Mytilus edulis*	97.885					239	301
Nemertea	Hoplonemertea	Monostilifera	*Ototyphlonemertes santacruzensis* (*)	81.481					107	123
Rotifera	Eurotatoria	Ploima	*Synchaeta triophthalma*	98.782			1945			

Table 7: Taxa detected in the August 2020 plankton samples via COI metabarcoding. Species marked with an asterisk are known to be introduced. Species marked with an asterisk in parentheses have not been recorded from the Belgian part of the North Sea yet.

Phylum	Class	Order	Name	Max. alignment (%)	Reads per taxon					
Location Mesh size					PO (100 µm)	PO (300 µm)	MM (100 µm)	MM (300 µm)	VD (100 µm)	VD (300 µm)
Annelida	Polychaeta	Spionida	*Boccardia pugettensis* (*)	84.743		197	722	3435	1081	6127
Arthropoda	Hexanauplia	Cyclopoida	*Oithona davisae* (*)	98.784					392	
	Malacostraca	Decapoda	*Ciliopagurus strigatus* (*)	100			287			
Cnidaria	Hydrozoa	Anthoathecata	*Corymorpha nutans*	83.077			2263	903		
		Leptothecata	*Cyclocanna producta*	81.579				259		
			Kirchenpaueria pinnata	81.164	1871	106	3105	2038		
			Obelia dichotoma	81.786					186	
		Siphonophorae	*Athorybia rosacea*	94.872					161	
Mollusca	Bivalvia	Ostreida	*Magallana gigas* *	99.538	451		101		157	

II.5: Overlap among different monitoring methods

Only nine taxa were identified to species level with high certainty (> 97 % alignment for metabarcoding data) via more than one method (Table 8). Of these, only *Ciona intestinalis* was found via more than two methods. Three of these species (*C. intestinalis*, *F. enigmaticus*, and *M. gigas*) are listed as NIS for the Belgian part of the North Sea by Verleye et al. (2020). *Molgula retortiformis* and *O. davisae* have not yet been recorded for this area (Horton et al., 2021).

Table 8: All species detected by more than one identification method (only species with an alignment greater than 97% are considered for the metabarcoding data). Species known to be introduced are marked with an asterisk, and those not recorded for the Belgian part of the North Sea are marked with an asterisk in parentheses.

Species	Morphology (plates)	18S rRNA (plates)	18S rRNA (plankton)	COI (plankton)
Botryllus schlosseri	x	x		
Centropages hamatus		x	x	
Ciona intestinalis *	x	x	x	
Ficopomatus enigmaticus *		x	x	
Lichomolgus canui		x	x	
Magallana gigas *			x	x
Molgula retortiformis (*)		x	x	
Obelia longissima	x		x	
Oithona davisae (*)			x	x

Discussion and conclusions

I. Overview of the results

A wide range of taxa, both indigenous and non-indigenous, were recovered through the combined application of metabarcoding and morphological examination of the settlement plate communities. The DNA-based method resulted in the discovery of 48 taxa, of which 25 were matched with high confidence (97 % or greater) to a particular species in the SILVA database. In contrast, the traditional method yielded 36 taxa (using a conservative estimate), with 16 of them identified to species level, from the settlement plate communities. The larger diversity detected with metabarcoding is in line with previous studies comparing the two methods (Brown et al., 2015; Zaiko et al., 2016; von Ammon et al., 2018). In this study, the two methods displayed very little overlap in the taxa detected, with *Ciona intestinalis, Ascidiella* sp. and *Botryllus schlosseri* being the only species and genera discovered with metabarcoding that were also found through visual inspection. When only species-level identifications are considered, only 5 % of the species were detected with both methods. The findings from this

study are in line with those from Obst et al. (2020), who found that only 4 to 8 % of species on settlement plates could be detected using both metabarcoding and visual examination.

When comparing the two methods, metabarcoding had numerous advantages over morphological examination, but also carried its own challenges. Metabarcoding enabled the detection of numerous taxa that had been entirely missed during morphological examination due to their small size or mode of life (such as parasites), and it was also able to resolve some taxa that morphological examination could assign only to broad groupings (such as the case with the nematodes). However, fifteen species from the settlement plates detected via visual examination were not detected via 18S rRNA metabarcoding of these samples. Another interesting demonstration of the limitations of DNA-based monitoring is that the 18S rRNA and COI markers detected only three genera (*Obelia*, *Ascidiella*, and *Magallana*) in common from the same set of plankton samples. Below, the advantages and pitfalls of metabarcoding, as noted in this study, will be discussed in greater detail.

II. Advantages of metabarcoding

II.1 Taxonomic resolution of difficult species

Metabarcoding enabled many taxa to be identified to a more specific taxonomic level than possible through examination of preserved, intact specimens. In addition, DNA-based methods allowed the discovery of several putative species that may have been missed during visual analysis due to similarity with known and/or native species.

The representatives of *Botrylloides* on the settlement plates were identified via the 18S rRNA marker to belong to the invasive alien *B. violaceus*. The serpulid worms, whose origin was likewise unclear during morphological examination, were assigned to *Ficopomatus enigmaticus*. This tubeworm builds reefs and may compete with native filter feeders for food, although the cold temperatures in the North Sea limit its proliferation so that infestations are generally not as bad as, for example, in the Mediterranean (Verleye et al., 2020). However, increases in mean sea water temperature due to climate change may result in increasing dominance of this serpulid in the future.

Rhodophytes were found on the plates from Ponton Overzet but could not be determined any further than the level of the phylum. 18S rRNA metabarcoding resolved these organisms to the species level, with one species, *Ceramium rubrum*, being native (Horton et al., 2021), and the other species, *Diapse ptilota*, being recorded from Australia only (Guiry & Guiry, 2021). However, the low alignment of reads assigned to *D. ptilota* (96.5 %) indicates that identification is questionable.

The results of the 18S rRNA metabarcoding suggest that the barnacles morphologically identified as the native *Amphibalanus improvisus* were in fact the recognized NIS *A. amphitrite*. The latter species is tropical to subtropical in origin but has established itself in many temperate coastal waters despite the low temperatures in these areas (Verleye et al., 2020). *Striatobalanus amaryllis* is also a tropical species represented in the 18S rRNA reads of the settlement plate communities, but it has not yet been recorded for the North Sea (Horton et al., 2021). Interesting is that *Austrobalanus imperator* was represented by a large proportion of reads from the plates at Ponton Overzet (Table 3), even for the underside of the upper plate, where no barnacles were visually noted (although microscopic larvae may have been present). The species was also found in the negative control, indicating possible cross-contamination among samples. *A. imperator* has never been recorded for the North Sea (Horton et al., 2021). Much like the confirmed NIS *Austrominius modestus* (which was never detected via 18S metabarcoding), *A. imperator* is an Oceanian species (Newman & Ross, 1976) and may thus survive the conditions in the Ostend harbour.

II.2. Detection of planktonic larval stages of benthic fauna

Many marine NIS, just as many marine species in general, exhibit planktonic larval stages, and thus plankton identification plays an important part in monitoring potentially invasive organisms (Duarte et al., 2021). However, planktonic organisms are notoriously difficult to identify visually due to their small size and the lack of discernible adult diagnostic characters displayed by larvae.

Magallana (formerly known as *Crassostrea*) was found to be represented by the ubiquitous *M. gigas* by both the COI and 18S rRNA markers (Table 8). This well-known, introduced oyster was detected exclusively in the 100 µm fractions taken in August, although 18S rRNA analysis found it only in the samples from Ponton Overzet while COI analysis found it at all three locations.

18S metabarcoding also enabled the detection of the shell-boring worm *Polydora haswelli* from Vuurtorendok and Marina Mercator at both time points (Tables 3 and 4). This organism was originally known only from southeast Australia and New Zealand, where it infests the shells of not only native Oceanian bivalves but also *Magallana gigas* (Read, 2010; Horton et al., 2021). In 2012 it was recorded from off Korea from specimens collected at the turn of the century, showing that its original distribution was broader than originally thought (Sato-Okoshi et al., 2012). *Magallana gigas* is common throughout the harbour of Ostend and may thus provide a suitable host for *P. haswelli*. This and related *Polydora* species are considered serious pests in the shellfish farming industry, as their burrowing damages and weakens the shells of their hosts and ultimately reduces marketability of the molluscs (Read, 2010; Sato-Okoshi et al., 2012). The potential effects of this spionid worm from a conservation standpoint

(e.g. on native oyster reefs) can only be speculated at this point. Therefore, the detection of *P. haswelli* in this study is a cause for concern. The two other annelid species detected with the 18S marker, *Ficopomatus enigmaticus* and *Pseudopolydora paucibranchiata* (Table 5), are NIS already documented for the North Sea. *F. enigmaticus* is, as explained above, an invasive reef builder. *P. paucibranchiata* is likewise an introduced tube builder, but rather than encrusting hard surfaces with calcareous tubes as *F. enigmaticus* does it builds silty tubes in soft sediments, sometimes in densities of up to several thousands of individuals per square metre (Radashevsky et al., 2020). It exhibits a wide tolerance to unfavourable conditions (such as pollution) (Radashevsky et al., 2020), potentially giving it an advantage in a heavily anthropogenized harbour such as the one sampled in this study. Originally from Asia, *P. paucibranchiata* has spread throughout the northern Atlantic, northwest Pacific, and North Sea regions, and has been recorded from Belgian waters (Radashevsky et al., 2020; Horton et al., 2021).

The COI method detected other species of annelid larvae than 18S rRNA (Tables 6 and 7). *Barrukia cristata*, *Boccardia pugettensis*, and *Lanice conchilega* were represented in the COI reads from the samples, but all with a low alignment (less than 85 %) to the database. *L. conchilega* is the common, indigenous sand mason worm, while *B. pugettensis* is native to British Columbia and *B. cristata* is from the Southern Ocean (Horton et al., 2021). Given the low alignment (below 85 % in both cases), the detection of the two non-native species should be viewed with caution. The reads corresponding to either sequence fall well outside the so-called barcoding gap and therefore cannot be reliably identified (Bucklin et al., 2011).

Only three species of ascidians were recorded from the plankton samples (Tables 4 and 6). Both methods found *Ascidiella* exclusively in the 100 µm fraction of Marina Mercator samples taken in June, and COI barcoding resolved the reads to the species level (*A. aspersa*, a native; Horton et al., 2021). However, the COI approach missed the other two ascidians. The *Molgula* reads were assigned by 18S rRNA metabarcoding to *M. retortiformis*, which was also detected on the settlement plates using the same marker gene.

II.3 Detection of endoparasites

Metabarcoding also retrieved sequences belonging to the native *Lichomolgus canui* (Tables 3 and 4), which is a parasitic copepod living inside various tunicates including the observed *Ciona intestinalis* and *Clavelina lepadiformis* (Hayward & Ryland, 2017). This unusual taxon illustrates a potential strength of metabarcoding over morphology, namely the ability to detect species living a cryptic lifestyle (such as parasites within their hosts) (Zaiko et al., 2015; von Ammon et al., 2018). In the settlement plates, sequences from *Lankesteria halocynthiae* were detected as well (Table 3). This organism is an apicomplexan parasite of tunicates, documented to infest the intestines of the Pacific species *Halocynthia auratum* (Rueckert et

al., 2015) and not yet recorded for the North Sea in the World Register of Marine Species (WoRMS) database (Horton et al., 2021). It is unlikely that the corresponding reads represent an introduced organism, though. Lecudinid apicomplexans such as *L. halocynthiae* are highly host-specific (Rueckert et al., 2015), and together with the low database match of the corresponding reads, this fact suggests that the species detected from the plates may well have been a native organism. This does not imply that the organism should go unnoticed, as introduced parasites have been shown to cause drastic losses in the biodiversity of areas they were introduced into (see, for example, the unicellular oyster parasite *Bonamia ostreae* as recounted by Verleye et al., 2020).

II.4 Detection of microscopic species

Although diatoms were not targeted in this study, they did show up as “by-catch” in both the settlement plate and plankton samples. These organisms are microscopic in size and, much like the larvae of benthic animals, difficult to identify morphologically. Nonetheless, at least three species of diatoms have been introduced into the Belgian part of the North Sea so far (Verleye et al., 2020). Of the diatoms found on the settlement plates (Table 3), all were previously recorded for the North Sea except for *Thalassiosira alleni*, which is recorded from the Indopacific region (Horton et al., 2021). The 18S rRNA marker analysis of plankton samples (Tables 4 and 5) turned up other species of diatoms, some of which (*Hemiaulus sinensis*, *Minidiscus* sp., *Skeletonema grethae*, and *Thalassiosira oceanica*) had not been recorded for the Belgian part of the North Sea in the World Register of Marine Species (Horton et al., 2021). *H. sinensis* and *T. oceanica* are fairly wide-ranging species and may simply not have been documented for the study region yet (both have been detected in the Baltic, for example). On the other hand, Gollasch et al. (2000) found that *Hemiaulus* and *Thalassiosira* species could survive long voyages (23 days in their study) in ships’ ballast water tanks, therefore representing a possible introduction. *Skeletonema grethae* was split off from the cosmopolitan species *S. costatum* in 2005 by Sarno and colleagues, who described the new species from the Pacific and Atlantic coasts of North America (Sarno et al., 2005). *Minidiscus*, which is one of the smallest diatoms in the world, is a cosmopolitan genus too (Fernandes & Correr-da-Silva, 2020).

Other unicellular organisms besides diatoms were found in the plankton and plate samples, including the bioluminescent dinoflagellate *Noctiluca scintillans* (Table 4) and the water mould *Haliphthoros* sp. (Table 3). The latter species is foreign to the study region (documented from North America; Horton et al., 2021), but its low read alignment may well indicate that the species from which the sequences came was simply a native species not yet in the SILVA database. The Nucletmycea sequences (described as a fungal metagenome by the database) and the Bicosoecida sequences (said to be from a freshwater species) also exhibited alignments lower than the commonly used threshold of 97 % (Table 3; Brown et al., 2015).

Therefore, the corresponding reads may well be of native marine organisms not yet recorded in the database.

Nematodes, ostracods, copepods, and platyhelminths, although multicellular organisms, are hardly visible to the naked eye. Although nematodes and copepods were noted in the settlement plates via visual identification, they generally could not be identified beyond phylum and subclass, respectively (except for a few specimens of the genus *Halicyclops*, which has a characteristic antennal morphology). 18S rRNA metabarcoding of settlement plate communities, however, revealed nine taxa of copepods and seven taxa of nematodes, with one copepod (*Paramphiascella fulvofasciata*) having previously been recorded only from littoral zones in Massachusetts Bay (Rosenfield & Coull, 1974). Ostracods and platyhelminths, which were never observed during visual analysis, were assigned to four different taxa, two of which (the ostracod *Sclerochilus oshoroensis* and the platyhelminth *Kaitalugia* sp.) were recorded only from Asia and Oceania, respectively (Horton et al., 2021).

In the plankton samples, 18S rRNA (but not COI) metabarcoding was able to detect the non-indigenous copepods *Acartia tonsa* and *Pseudodiaptomus marinus* (Table 5), which are listed NIS in the work of Verleye et al. (2020). These species can become locally abundant due to their broad salinity and temperature ranges; in addition, *A. tonsa* survives unfavourable conditions via the production of resting eggs. 18S rRNA metabarcoding also found two copepod taxa (*Oithona* sp. and *Pseudanthessius* sp.; Tables 3 and 4) described in the database as being from New Caledonia; these taxa may represent a potential introduction but cannot be definitively described as such since the species names were never given for these reads. Another non-indigenous copepod, *Oithona davisae*, was found both via 18S rRNA and via COI metabarcoding (Table 8). This species has never been recorded from the Belgian part of the North Sea yet (Horton et al., 2021), but Cornils & Wend-Heckmann (2015) confirmed its presence in the Wadden Sea. Here, it was reported to reach high abundances, sometimes becoming the most prevalent copepod species after members of *Acartia*. *Oithona davisae* is Northwest Pacific in origin and known to be invasive in the Black and northwest Mediterranean Seas. As the original habitat of *O. davisae* is generally warmer than the Wadden Sea and other parts of the North Sea, climate change may play a role in the proliferation of this species in its new range (Cornils & Wend-Heckmann, 2015).

III. Challenges of the metabarcoding approach

III.1: Low taxonomic resolution of barcoding loci

Even though many more taxa were assigned to species level with metabarcoding than with morphology, the assignment of a group of reads to a taxon—whether native or non-native—should not be accepted unquestioningly. Multiple closely related species may share a (nearly)

identical COI or 18S rRNA marker region, making their separation impossible or nearly so. This problem is especially important in the detection of NIS, which may be closely related to native species (Duarte et al., 2021).

For example, the *Molgula* on the plates, which had been suspected to be non-native (possibly the well-documented *M. manhattensis*; Verleye et al., 2020), produced sequences assigned to foreign species (*M. provisionalis* and *M. retortiformis*; Table 3) never documented for the North Sea. Much like *M. manhattensis*, *M. provisionalis* and *M. retortiformis* are originally from the Northwest Atlantic region (Verleye et al., 2020; Horton et al., 2021). Although they may have established themselves in this area following transport via shipping traffic, they might also represent misassignment of the 18S rRNA marker. *Ascidia ceratodes, Symplegma viride* and *Styela plicata* went entirely unnoticed during morphological examination but were detected via the 18S rRNA marker (Table 3). These species have never been documented from the North Sea either; the first is native to the Northeast Pacific, the second has been recorded from the southern hemisphere and is introduced into the Mediterranean, while the third has a circumglobal distribution and is considered invasive in various areas in the Mediterranean (Horton et al., 2021). It is not unfeasible that the environment sampled in this study may provide them with a habitat sufficiently similar to that from which they were transported. However, the low read match of the generated sequences with those in the SILVA database for *S. viride* implies that this species may be in fact a misidentification. *Styela plicata* may well be the congeneric *Styela clava*, which has been recorded from the Belgian part of the North Sea (Verleye et al., 2020).

The difficulties of the 18S rRNA gene region in discriminating some species stems from the fact that it has a slower rate of mutation (relative to the COI gene), which leads to smaller differences between closely related species (Bucklin et al., 2011) and a tendency to underestimate biodiversity (van der Loos & Nijland, 2020). As explained in the introduction, the 18S marker has a broad taxonomic coverage that comes at the cost of decreased discriminatory ability for some taxa (Creer et al., 2016). For this reason, COI metabarcoding is preferred for macrobenthos communities (van der Loos & Nijland, 2020). However, in this study, COI-based analysis of the settlement plate communities was not possible due to technical difficulties. Had it been able to do so, COI metabarcoding may have yielded some or all of these taxa. It is noteworthy to mention that in the plankton samples, COI metabarcoding resolved the *Ascidiella* reads to species level, while 18S rRNA metabarcoding assigned them only to genus level.

Several COI-based plankton identifications, however, were also questionable. One possible instance is the hermit crab *Ciliopagurus strigatus* (Table 7). Although the 100% match with the database would appear to indicate an introduction (for example, of larvae via ballast water), it is theoretically possible that the COI sequence of this hermit crab is identical to that of native hermit crabs, leading to a misidentification. According to Hebert et al. (2003), this is unlikely.

These authors found that, in a set of over 1780 congeneric crustacean species pairs, less than 5 % of these pairs exhibited sequence divergence of below 4 %. Similar results were obtained for most animal phyla surveyed by these authors, except for cnidarians, in which 88.2 % of congeneric species pairs exhibited less than 1 % divergence.

III.2: Completeness and correctness of reference libraries

Fifteen species detected via visual examination of settlement plate communities were not found via 18S rRNA metabarcoding. Several explanations exist for the lack of coverage of these organisms. First, they may be absent from the SILVA database. Manually searching this database (https://www.arb-silva.de/search/) revealed no entries for *Abludomelita obtusata*, *Anoplodactylus virescens* (although the genus *Anoplodactylus* was represented), *Crassicorophium bonellii*, *Maera grossimana*, *Microdeutopus gryllotalpa*, and *Pedicellina hispida*. Absence of a species from a database can also lead to its reads being misidentified rather than unidentified, leading to false positives (Zaiko et al., 2016). In this study, the prevalence of taxa never recorded for the Belgian coast may imply that some of these putative taxa are such false positives. Incompleteness of reference libraries is currently a major problem of DNA-based biodiversity surveys (Zaiko et al., 2015; Duarte et al., 2021). Even though the North Sea is one of the best-studied marine regions in the world according to Hestetun et al. (2020), these authors found that only 42.4 % and 27.1 % of macrobenthos species detected morphologically near Norwegian oil platforms were represented in the BOLD and SILVA databases, respectively.

Not only are missing entries a problem in databases, but also the fact that some reference sequences originate from misidentified organisms (Kress et al., 2015) or contain sequencing errors (Valentini et al., 2009). In this study, for example, the putative *Austrobalanus imperator* (Table 3) may have come from misidentified *Austrominius modestus*: both species share the same native range and are in the same family, suggesting that *A. modestus*, visually detected on the settlement plates, may have been attributed to reads that really did originate from this species. The same can be said for the assignment of *Molgula retortiformis* to reads from both the settlement plate and plankton samples, as the species originates from the same area as the known NIS *M. manhattensis*. In this respect metabarcoding is also dependent to some extent on taxonomic expertise, with possible errors amplifying across studies rather than being limited to the study in which the misidentification occurred.

III.3: Technical false negatives

Some morphologically detected groups on the settlement plates were not found by metabarcoding even though they were represented in the SILVA database. These taxa include some ascidians (namely *Clavelina lepadiformis* and *Diplosoma listerianum*), skeleton shrimps

(*Caprella* sp.), Stauromedusae, syllid worms, etc. Two confirmed non-indigenous taxa—the barnacle *Austrominius modestus* and the crab family Varunidae—were never found via metabarcoding, despite being included in the SILVA database. *A. modestus* is an important fouling species, whose rapid growth allows it to compete effectively with native and commercially important sessile organisms (Verleye et al., 2020). These authors mention several species of Varunidae that are well-established in the Belgian part of the North Sea and implicated in competing with native crabs. The failure of metabarcoding to detect these species points to the fact that DNA-based detection methods are not immune to generating false negatives (Duarte et al., 2021) despite their potential for high sensitivity and accuracy (von Ammon et al., 2018).

Notably, many of these species were small organisms represented by only a few individuals per sample. The low biomass of many of these organisms in the samples may have led to the exclusion of their genetic material from the 0.25 g subsamples taken for DNA extraction, despite efforts to homogenize the samples as thoroughly as possible beforehand. In order to avoid such biomass-dependent biases, van der Loos & Nijland (2020) recommend sorting samples according to size in addition to carrying out a homogenization. Another possibility is primer mismatch with the DNA sequences of these organisms, as well as amplification bias in general (Trebitz et al., 2017; Duarte et al., 2021).

As explained in the materials and methods section, the two most speciose plate samples had three instead of one DNA extraction performed on them to maximize the probability of detecting rare species. One of these subsamples (the second replicate of PO-LoUp) was sequenced in the first run. As can be seen from the sequencing output, a great disparity exists between the two subsamples (Table 3). One would expect that most species in the first subsample would be found in the second, but this is hardly the case as only two of the ten taxa from the same side of the settlement plate were shared between both subsamples. Six taxa were found exclusively in one of the two subsamples and not from any other settlement plate sides; some of these taxa would have gone unnoticed if only one PO-LoUp replicate was used. The disagreement in recorded taxa from the two subsamples is, again, likely due to amplification bias or limited read coverage in one of the replicates. Due to multiple causes, such as differences in primer affinity or the presence of other biological material, DNA is not always representatively extracted and amplified from a sample, so that some species may be underrepresented or even absent in the final analysis (Trebitz et al., 2017). In light of this possibility, the results from the other samples must be treated with caution, as metabarcoding may too have missed many species from these samples. A further improvement would therefore be to sequence multiple subsamples from all the settlement plate sides in order to increase taxonomic coverage. Greater sequence depth can also offset biases due to PCR amplification and preferential primer annealing, but it does this at the cost of increased number of sequence artefacts (van der Loos & Nijland, 2020).

III.4: Lack of ability to estimate relative abundances

The majority of settlement plate-derived reads across all samples belonged to arthropods and chordates (Figure 7), possibly due to the large contribution of barnacles and tunicates, respectively, to the biomass on the fouling plates. As explained above, species with a low abundance and/or biomass tend to be less represented in metabarcoding analyses of communities, hence the recommendation of van der Loos & Nijland (2020) to include size sorting and homogenization steps. However, these authors also caution that proportion of reads is often weakly or not at all correlated with the proportion of species found in the actual community. Therefore, metabarcoding cannot be used to accurately estimate abundances of organisms in their natural settings. A morphological examination, on the other hand, might miss many organisms but can give a reasonably accurate estimate of the relative abundances of the most dominant species in a community. Such estimates of abundance may be useful, for example, to determine if an eradication plan for an established NIS still is feasible.

IV. Evaluation of the performance of the metabarcoding approach

This study demonstrated the capacity of metabarcoding to survey and identify a much larger array of species than would be possible using traditional, morphology-based methods. Some of the species detected via metabarcoding were not typical North Sea species (native or confirmed introduced species) and may represent potential introductions. If this is indeed the case, ballast water transport and/or hull fouling would form the most likely vectors for the transport of these species, as none of them are of commercial interest and most of them are planktonic during at least some part of their life stages.

However, the great disparity between morphological and metabarcoding methods in terms of taxa detected is a well-known point of concern (van der Loos & Nijland, 2020). Had the metabarcoding method been accurate, one would have expected it to have turned up most or all of the visually identified species in the settlement plate samples. Instead, only two species were detected from these samples via both methods. Given this disparity, it is likely that some or even most of the species detected for the first time in the North Sea are simply false positives. Especially the 18S rRNA data must be viewed with caution, as this marker is generally not suited for species resolution in many taxa (Tang et al., 2012; Creer et al., 2016).

Given the potential inaccuracy of DNA-based detection methods, one conclusion from this pilot study is the importance of benchmarking such results against an inventory constructed based on morphological data. It was only because a morphological examination of settlement plate samples preceded their metabarcoding that the issues with the latter method became apparent. Plankton samples were not identified morphologically, but the use of two different marker genes yielded quite divergent results, highlighting the same concerns. Therefore, until

further improvements are made in the accuracy and capacity of metabarcoding, future NIS surveys should ideally include morphological as well as DNA-based monitoring methods. Multiple marker genes should be used, especially for diverse communities like those analysed in this study. The use of more than one detection method is recommended by numerous authors, such as Zaiko et al. (2015), Trebitz et al. (2017), and von Ammon et al. (2018). However, Creer et al. (2016) caution that the different strengths of individual primers do not necessarily make up for the other primers' weaknesses. In addition, employing more taxon-specific primers in targeted surveys may help to ameliorate the issue of low primer complementarity (Trebitz et al., 2017).

Even within a sample, metabarcoding yielded highly different estimates of species richness, as demonstrated by the two subsamples taken from PO-LoUp. It can therefore be concluded that increasing sampling size would increase the detection ability of the metabarcoding approach, and possibly among-sample consensus as well. Ideally, a collector's curve must be constructed for the number of species detected with increasing sampling effort to assess the thoroughness of the survey, as was done by Zaiko et al. (2016).

One further method of accounting for the bias of metabarcoding may be to include among the samples a positive control, namely a mock community consisting of known species in known proportions (Holman et al., 2019; Duarte et al., 2021). For example, Brown et al. (2015) used mock communities to optimize taxon-specific divergence thresholds when metabarcoding with the 18S V4 region. Zaiko et al. (2016), when metabarcoding settlement plate communities with the 18S rRNA marker gene, employed a positive control of mussel, polychaete, and ascidian DNA mixed in various proportions. These authors used this control to demonstrate differences in taxon resolution achieved by the 18S rRNA method as well as the correlation between read number and amount of original DNA.

Finally, the use of a location-specific reference library, constructed with species known to occur only in the area of study, offers a potential means of avoiding misidentifications (Zaiko et al., 2015; Zaiko et al., 2016; Holman et al., 2019). To this end, the GEANS project collects specimens identified to species level into voucher collections, which will then be used to construct a reference database specific for the North Sea (https://northsearegion.eu/geans/about/). However, it should be noted that exclusive use of a location-specific reference database may miss species newly introduced into the area and not yet entered into the database. The best approach in this case may be to consult a location-specific reference database alongside a more general one like SILVA or MZGdb.

V. Evaluation of the performance of the MinION in DNA-based NIS surveys

A serious shortcoming of the MinION is the high error rate (10 – 22%) of raw reads (Baloğlu et al., 2021). Variability in the activity of enzymes that escort the DNA molecules through the

pores contributes to these errors (Deamer et al., 2016). In addition, and more importantly, the error rate is increased due to the influence that neighbouring nucleotides exert on the current drop initiated by the nucleotide in the pore (Deamer et al., 2016). The median Q scores (Phred quality scores) of both MinION runs in this study stayed around 8 for the first run and 10 for the second (figures S5 and S6), indicating an error rate of 17 % and 10%, respectively. These error rates are in line with those reported earlier for this device. The high error rates specific to the MinION, however, can be ameliorated by clustering sequences and forming consensus sequences (Baloğlu et al., 2021), which were also done in this study. Loit et al. (2019) found that this device was able to effectively detect dominant species in a fungal community, but it missed many of the less abundant species. These authors note that the ability of the MinION to sequence longer barcodes may partly offset this observed deficiency. In this study, however, the typically used short barcodes were sequenced.

Nonetheless, the present study benefitted from the MinION. Had the amplified DNA been sequenced using second-generation methods, the study time would have been much prolonged, especially if the amplicons were sent to a commercial sequencing company for analysis. The increased speed with which results could be obtained was in fact one of the main motivations for employing this device.

VI. Conclusions and future perspectives

Metabarcoding provides a new dimension of information that traditional biodiversity surveys have never achieved. Where human taxonomists struggle to identify small, cryptic, and/or poorly-known organisms, which includes a large portion of the world's documented species, DNA sequencing can rapidly assign a name to an unknown specimen from a complex community. In many cases, this assignment is fairly accurate. In other cases, the community composition as reported by the read output is dubious. Every step of the metabarcoding process, from experimental design to choice of primers to bioinformatic parameters, has the potential to increase or decrease the occurrence of such errors.

With these limitations in mind, the best approach to monitoring NIS at the present seems to involve a combination of traditional and DNA-based methods, so that the benefits of either method make up (as much as possible) for the other method's shortcomings. Also, different approaches to metabarcoding of NIS should be attempted and compared, including more specific primers, location-specific reference databases, positive controls, and larger sample sizes. Different sequencing technologies should also be compared, including less traditional and more novel ones such as the MinION. In this case, the MinION may not have produced highly accurate reads, but its small size and shorter workflow were advantageous relative to other sequencers. A thorough study of different approaches to metabarcoding will aid the timely and accurate detection of NIS in harbours.

References

Anton, A., Geraldi, N. R., Lovelock, C. E., Apostolaki, E. T., Bennett, S., Cebrian, J., Krause-Jensen, D., Marbà, N., Martinetto, P., Pandolfi, J. M., Santana-Garcon, J., & Duarte, C. M. (2019). Global ecological impacts of marine exotic species. *Nature Ecology & Evolution*, *3*(5), 787-800. https://doi.org/10.1038/s41559-019-0851-0

Atalah, J., Newcombe, E. M., & Zaiko, A. (2016). Biocontrol of fouling pests: effect of diversity, identity and density of control agents. *Marine Environmental Research*, *115*, 20-27. https://doi.org/10.1016/j.marenvres.2016.01.006

Atalah, J., Newcombe, E. M., Hopkins, G. A., & Forrest, B. M. (2014). Potential biocontrol agents for biofouling on artificial structures. *Biofouling*, *30*(8), 999-1010. https://doi.org/10.1080/08927014.2014.956734

Balaji, R., Yaakob, O., & Koh, K. K. (2014). A review of developments in ballast water management. *Environmental Reviews*, *22*(3), 298-310. https://doi.org/10.1139/er-2013-0073

Baloğlu, B., Chen, Z., Elbrecht, V., Braukmann, T., MacDonald, S., & Steinke, D. (2021). A workflow for accurate metabarcoding using nanopore MinION sequencing. *Methods in Ecology and Evolution, 12*, 794-804. https://doi.org/10.1111/2041-210X.13561

Bauer, J. T. (2012). Invasive species: “back-seat drivers” of ecosystem change?. *Biological Invasions*, *14*(7), 1295-1304. https://doi.org/10.1007/s10530-011-0165-x

Bishop, M. J., & Hutchings, P. A. (2011). How useful are port surveys focused on target pest identification for exotic species management?. *Marine Pollution Bulletin*, *62*(1), 36-42. https://doi.org/10.1016/j.marpolbul.2010.09.014

Bott, N.J. (2015). *Detection of Invasive Species*. In: Kim S. K. (eds), *Springer Handbook of Marine Biotechnology* (pp. 441-453). Springer. https://doi.org/10.1007/978-3-642-53971-8_16

Boudouresque, C. F., & Verlaque, M. (2012). An overview of species introduction and invasion processes in marine and coastal lagoon habitats. *CBM-Cahiers de Biologie Marine*, *53*(3), 309-317.

Bray, J.R. & J.T. Curtis. 1957. An ordination of the upland forest communities of Southern Wisconsin. *Ecological Monographs*, *27*(4), 325-349. https://doi.org/10.2307/1942268

Brown, E. A., Chain, F. J., Crease, T. J., MacIsaac, H. J., & Cristescu, M. E. (2015). Divergence thresholds and divergent biodiversity estimates: can metabarcoding reliably

describe zooplankton communities?. *Ecology and Evolution*, *5*(11), 2234-2251. https://doi.org/10.1002/ece3.1485

Brown, E. A., Chain, F. J., Zhan, A., MacIsaac, H. J., & Cristescu, M. E. (2016). Early detection of aquatic invaders using metabarcoding reveals a high number of non-indigenous species in Canadian ports. *Diversity and Distributions*, *22*(10), 1045-1059. https://doi.org/10.1111/ddi.12465

Buckley, Y. M., & Catford, J. (2016). Does the biogeographic origin of species matter? Ecological effects of native and non-native species and the use of origin to guide management. *Journal of Ecology*, *104*(1), 4-17. https://doi-org.kuleuven.ezproxy.kuleuven.be/10.1111/1365-2745.12501

Bucklin, A., Peijnenburg, K. T., Kosobokova, K. N., O'Brien, T. D., Blanco-Bercial, L., Cornils, A., Falkenhaug, T., Hopcroft, R. R., Hosia, A., Laakmann, S., Li, C., Martell, L., Questel., J. M., Wall-Palmer, D., Wang, M., Wiebe, P. H., & Weydmann-Zwolicka, A. (2021). Toward a global reference database of COI barcodes for marine zooplankton. *Marine Biology*, *168*(6), 1-26. https://doi.org/10.1007/s00227-021-03887-y

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

Campbell, M. L., Gould, B., & Hewitt, C. L. (2007). Survey evaluations to assess marine bioinvasions. *Marine Pollution Bulletin*, *55*(7-9), 360-378. https://doi.org/10.1016/j.marpolbul.2007.01.015

Cardeccia, A., Marchini, A., Occhipinti-Ambrogi, A., Galil, B., Gollasch, S., Minchin, D., Narščius, A., Olenin, S., & Ojaveer, H. (2018). Assessing biological invasions in European Seas: Biological traits of the most widespread non-indigenous species. *Estuarine, Coastal and Shelf Science*, *201*, 17-28. https://doi.org/10.1016/j.ecss.2016.02.014

Casas-Monroy, O., Linley, R. D., Adams, J. K., Chan, F. T., Drake, D. A. R., & Bailey, S. A. (2015). Relative invasion risk for plankton across marine and freshwater systems: examining efficacy of proposed international ballast water discharge standards. *PloS One*, *10*(3), e0118267. https://doi.org/10.1371/journal.pone.0118267

Cheung, W. W., Lam, V. W., Sarmiento, J. L., Kearney, K., Watson, R., & Pauly, D. (2009). Projecting global marine biodiversity impacts under climate change scenarios. *Fish and Fisheries*, *10*(3), 235-251. https://doi.org/10.1111/j.1467-2979.2008.00315.x

Cornils, A., & Wend-Heckmann, B. (2015). First report of the planktonic copepod Oithona davisae in the northern Wadden Sea (North Sea): Evidence for recent invasion?. *Helgoland Marine Research*, *69*(2), 243-248. https://doi.org/10.1007/s10152-015-0426-7

Corrales, X., Katsanevakis, S., Coll, M., Heymans, J. J., Piroddi, C., Ofir, E., & Gal, G. (2020). Advances and challenges in modelling the impacts of invasive alien species on aquatic ecosystems. *Biological Invasions, 22*, 907-934. https://doi.org/10.1007/s10530-019-02160-0

Creer, S., Deiner, K., Frey, S., Porazinska, D., Taberlet, P., Thomas, W. K., Potter, C., & Bik, H. M. (2016). The ecologist's field guide to sequence-based identification of biodiversity. *Methods in Ecology and Evolution*, *7*(9), 1008-1018. https://doi.org/10.1111/2041-210X.12574

Cristescu, M. E., & Hebert, P. D. (2018). Uses and misuses of environmental DNA in biodiversity science and conservation. *Annual Review of Ecology, Evolution, and Systematics*, *49*(1), 209-230.

Cullings, K. W. (1992). Design and testing of a plant-specific PCR primer for ecological and evolutionary studies. *Molecular Ecology*, *1*(4), 233-240. https://doi.org/10.1111/j.1365-294X.1992.tb00182.x

D'hondt, B., Vanderhoeven, S., Roelandt, S., Mayer, F., Versteirt, V., Adriaens, T., Ducheyne, E., San Martin, G., Grégoire, J., Stiers, I., Quoilin, S., Cigar, J., Heughebaert, A., & Branquart, E. (2015). Harmonia+ and Pandora+: risk screening tools for potentially invasive plants, animals and their pathogens. *Biological Invasions, 17*(6), 1869-1883. https://doi.org/10.1007/s10530-015-0843-1

Dafforn, K. A., Johnston, E. L., & Glasby, T. M. (2009). Shallow moving structures promote marine invader dominance. *Biofouling, 25*(3), 277-287. https://doi.org/10.1080/08927010802710618

Deamer, D., Akeson, M., & Branton, D. (2016). Three decades of nanopore sequencing. *Nature Biotechnology, 34*(5), 518-524. https://doi.org/10.1038/nbt.3423

Didham, R. K., Tylianakis, J. M., Hutchison, M. A., Ewers, R. M., & Gemmell, N. J. (2005). Are invasive species the drivers of ecological change?. *Trends in Ecology & Evolution, 20*(9), 470-474. https://doi.org/10.1016/j.tree.2005.07.006

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

Essl, F., Bacher, S., Blackburn, T. M., Booy, O., Brundu, G., Brunel, S., Cardoso, A., Eschen, R., Gallardo, B., Galil, B., García-Berthou, E., Genovesi, P., Groom, Q., Harrower, C., Hulme, P. E., Katsanevakis, S., Kenis, M., Kühn, I., Kumschik, S., … & Jeschke, J. M. (2015). Crossing frontiers in tackling pathways of biological invasions. *BioScience, 65*(8), 769-782. https://doi.org/10.1093/biosci/biv082

Essl, F., Bacher, S., Genovesi, P., Hulme, P. E., Jeschke, J. M., Katsanevakis, S., Kowarik, I., Kühn, I., Pyšek, P., Rabitsch, W., Schindler, S., van Kleunen, M., Vilà, M., Wilson, J. R. U., & Richardson, D. M. (2018). Which taxa are alien? Criteria, applications, and uncertainties. *BioScience, 68*(7), 496-509. https://doi.org/10.1093/biosci/biy057

European Commission. (2020, July). The Marine Strategy Framework Directive. European Commission. https://ec.europa.eu/environment/marine/eu-coast-and-marine-policy/marine-strategy-framework-directive/index_en.htm

Faulkner, K. T., Robertson, M. P., Rouget, M., & Wilson, J. R. (2014). A simple, rapid methodology for developing invasive species watch lists. *Biological Conservation, 179*, 25-32. https://doi.org/10.1016/j.biocon.2014.08.014

Faust, E., André, C., Meurling, S., Kochmann, J., Christiansen, H., Jensen, L. F., Charrier, G., Laugen, A. T., & Strand, Å. (2017). Origin and route of establishment of the invasive Pacific oyster Crassostrea gigas in Scandinavia. *Marine Ecology Progress Series*, 575, 95-105. https://doi.org/10.3354/meps12219

Fernandes, L. F., & Correr-Da-Silva, F. (2020). Diversity and distribution of nanoplanktonic Minidiscus (Bacillariophyta) in shelf waters of the Southwest Atlantic Ocean. *Diatom Research, 35*(4), 327-338. https://doi.org/10.1080/0269249X.2020.1845807

Fraija-Fernández, N., Bouquieaux, M. C., Rey, A., Mendibil, I., Cotano, U., Irigoien, X., Santos, M., & Rodríguez-Ezpeleta, N. (2020). Marine water environmental DNA metabarcoding provides a comprehensive fish diversity assessment and reveals spatial patterns in a large oceanic area. *Ecology and Evolution, 10*(14), 7560-7584. https://doi.org/10.1002/ece3.6482

Fridley, J. D., & Sax, D. F. (2014). The imbalance of nature: revisiting a Darwinian framework for invasion biology. *Global Ecology and Biogeography, 23*(11), 1157-1166. https://doi.org/10.1111/geb.12221

García-Berthou, E., Alcaraz, C., Pou-Rovira, Q., Zamora, L., Coenders, G., & Feo, C. (2005). Introduction pathways and establishment rates of invasive aquatic species in Europe. *Canadian Journal of Fisheries and Aquatic Sciences, 62*(2), 453-463. https://doi.org/10.1139/f05-017

Geller, J., Meyer, C., Parker, M., & Hawk, H. (2013). Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Molecular Ecology Resources, 13*(5), 851-861. https://doi.org/10.1111/1755-0998.12138

Gilroy, J. J., Avery, J. D., & Lockwood, J. L. (2017). Seeking international agreement on what it means to be "native". *Conservation Letters, 10*(2), 238-247. https://doi.org/10.1111/conl.12246

Gollasch, S., Lenz, J., Dammer, M., & Andres, H. G. (2000). Survival of tropical ballast water organisms during a cruise from the Indian Ocean to the North Sea. *Journal of Plankton Research, 22*(5), 923-937. https://doi.org/10.1093/plankt/22.5.923

Grosholz, E. D., Crafton, R. E., Fontana, R. E., Pasari, J. R., Williams, S. L., & Zabin, C. J. (2015). Aquaculture as a vector for marine invasions in California. *Biological Invasions, 17*(5), 1471-1484. https://doi.org/10.1007/s10530-014-0808-9

Guiry, M.D. & Guiry, G.M. (2021). Diapse ptilota (J.D.Hooker & Harvey) Kylin 1956. AlgaeBase. World-wide electronic publication, National University of Ireland, Galway. http://www.algaebase.org

Guy-Haim, T., Hyams-Kaphzan, O., Yeruham, E., Almogi-Labin, A., & Carlton, J. T. (2017). A novel marine bioinvasion vector: Ichthyochory, live passage through fish. *Limnology and Oceanography Letters, 2*(3), 81-90. https://doi.org/10.1002/lol2.10039

Hadziavdic, K., Lekang, K., Lanzen, A., Jonassen, I., Thompson, E. M., & Troedsson, C. (2014). Characterization of the 18S rRNA gene for designing universal eukaryote specific primers. *PloS One, 9*(2), e87624. https://doi.org/10.1371/journal.pone.0087624

Hayward, P. J., & Ryland, J. S. (Eds.). (2017). *Handbook of the Marine Fauna of North-West Europe* (2nd ed.). Oxford University Press.

Hebert, P. D., Ratnasingham, S., & De Waard, J. R. (2003). Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 270*(suppl_1), S96-S99. https://doi.org/10.1098/rsbl.2003.0025

Hestetun, J. T., Bye-Ingebrigtsen, E., Nilsson, R. H., Glover, A. G., Johansen, P. O., & Dahlgren, T. G. (2020). Significant taxon sampling gaps in DNA databases limit the operational use of marine macrofauna metabarcoding. *Marine Biodiversity, 50*(5), 1-9. https://doi.org/10.1007/s12526-020-01093-5

Hewitt, C. L., & Martin, R. B. (2001). Revised protocols for baseline port surveys for introduced marine species–design considerations, sampling protocols and taxonomic sufficiency. *Centre for Research on Introduced Marine Pests Technical Report*, (22).

Holman, L. E., de Bruyn, M., Creer, S., Carvalho, G., Robidart, J., & Rius, M. (2019). Detection of introduced and resident marine species using environmental DNA metabarcoding of sediment and water. *Scientific Reports, 9*(1). https://doi.org/10.1038/s41598-019-47899-7

Hopkins, G. A., Prince, M., Cahill, P. L., Fletcher, L. M., & Atalah, J. (2016). Desiccation as a mitigation tool to manage biofouling risks: trials on temperate taxa to elucidate factors influencing mortality rates. *Biofouling, 32*(1), 1-11. https://doi.org/10.1080/08927014.2015.1115484

Horton, T., Kroh, A., Ahyong, S., Bailly, N., Boyko, C. B., Brandão, S. N., Gofas, S., Hooper, J. N. A., Hernandez, F., Holovachov, O., Mees, J., Molodtsova, T. N., Paulay, G., Decock, W., Dekeyzer, S., Poffyn, G., Vandepitte, L., Vanhoorne, B., Adlard, R., ... & Zhao, Z. (2021). World Register of Marine Species. https://www.marinespecies.org

Hulme, P. E. (2009). Trade, transport and trouble: managing invasive species pathways in an era of globalization. *Journal of Applied Ecology, 46*(1), 10-18. https://doi.org/10.1111/j.1365-2664.2008.01600.x

Hulme, P. E., Bacher, S., Kenis, M., Klotz, S., Kühn, I., Minchin, D., Nentwig, W., Olenin, S., Panov, V., Pergl, J., Pyšek, P., Roques, A., Sol, D., Solarz, W., & Vilà, M. (2008). Grasping at the routes of biological invasions: a framework for integrating pathways into policy. *Journal of Applied Ecology, 45*(2), 403-414. https://doi.org/10.1111/j.1365-2664.2007.01442.x

Jeschke, J. M. (2014). General hypotheses in invasion ecology. *Diversity and Distributions, 20*(11), 1229-1234. https://doi.org/10.1111/ddi.12258

Joyce, P. W., Cuthbert, R. N., Kregting, L., Crane, K., Vong, G. Y., Cunningham, E. M., Dick, J. T. A., & Coughlan, N. E. (2019). Stay clean: direct steam exposure to manage biofouling risks. *Marine Pollution Bulletin, 142*, 465-469. https://doi.org/10.1016/j.marpolbul.2019.04.011

Karatayev, A. Y., Burlakova, L. E., Padilla, D. K., Mastitsky, S. E., & Olenin, S. (2009). Invaders are not a random selection of species. *Biological Invasions, 11*(9). https://doi.org/10.1007/s10530-009-9498-0

Kettunen, M., Genovesi, P., Gollasch, S., Pagad, S., Starfinger, U. ten Brink, P. & Shine, C. 2009. *Technical support to EU strategy on invasive species (IAS) - Assessment of the impacts of IAS in Europe and the EU (final module report for the European Commission)*. Institute for

European Environmental Policy (IEEP), Brussels, Belgium. 44 pp. + Annexes.

Kono, N., & Arakawa, K. (2019). Nanopore sequencing: review of potential applications in functional genomics. *Development, Growth & Differentiation, 61*(5), 316-326. https://doi.org/10.1111/dgd.12608

Kress, W. J., García-Robledo, C., Uriarte, M., & Erickson, D. L. (2015). DNA barcodes for ecology, evolution, and conservation. *Trends in Ecology & Evolution, 30*(1), 25-35. https://doi.org/10.1016/j.tree.2014.10.008

Lane, S. E., Hollings, T., Hayes, K. R., McEnnulty, F. R., Green, M., Georgiades, E., & Robinson, A. P. (2018). Risk factors for fouling biomass: evidence from small vessels in Australia. *Biofouling, 34*(9), 1032-1045. https://doi.org/10.1080/08927014.2018.1536202

Lobo, J., Costa, P. M., Teixeira, M. A., Ferreira, M. S., Costa, M. H., & Costa, F. O. (2013). Enhanced primers for amplification of DNA barcodes from a broad range of marine metazoans. *BMC Ecology, 13*(1), 1-8. https://doi.org/10.1186/1472-6785-13-34

Lockwood, J. L., Cassey, P., & Blackburn, T. (2005). The role of propagule pressure in explaining species invasions. *Trends in Ecology & Evolution, 20*(5), 223-228. https://doi.org/10.1016/j.tree.2005.02.004

Loit, K., Kalev Adamson, M. B., Puusepp, R., Anslan, S., Kiiker, R., Drenkhan, R., & Tedersoo, L. (2019). Relative performance of MinION (Oxford Nanopore Technologies) versus Sequel (Pacific Biosciences) third-generation sequencing instruments in identification of agricultural and forest fungal pathogens. *Applied and Environmental Microbiology, 85*(21). https://dx.doi.org/10.1128%2FAEM.01368-19

Lyons, T. J., Tuckett, Q. M., Donahou, A. D., & Hill, J. E. (2020). Risk screen of lionfishes, Pterois, Dendrochirus, and Parapterois, for southeastern United States coastal waters of the Gulf of Mexico and Atlantic Ocean. *Biological Invasions, 22*, 1573-1583. https://doi-org.kuleuven.ezproxy.kuleuven.be/10.1007/s10530-020-02203-x

McGeoch, M. A., Genovesi, P., Bellingham, P. J., Costello, M. J., McGrannachan, C., & Sheppard, A. (2016). Prioritizing species, pathways, and sites to achieve conservation targets for biological invasion. *Biological Invasions, 18*(2), 299-314. http://dx.doi.org.kuleuven.ezproxy.kuleuven.be/10.1007/s10530-015-1013-1

Meyer, C. P., Geller, J. B., & Paulay, G. (2005). Fine scale endemism on coral reefs: archipelagic differentiation in turbinid gastropods. *Evolution, 59*(1), 113-125. https://doi.org/10.1111/j.0014-3820.2005.tb00899.x

Molnar, J. L., Gamboa, R. L., Revenga, C., & Spalding, M. D. (2008). Assessing the global threat of invasive species to marine biodiversity. *Frontiers in Ecology and the Environment, 6*(9), 485-492. https://doi.org/10.1890/070064

Montefalcone, M., Morri, C., Parravicini, V., & Bianchi, C. N. (2015). A tale of two invaders: divergent spreading kinetics of the alien green algae Caulerpa taxifolia and Caulerpa cylindracea. *Biological Invasions, 17*(9), 2717-2728. https://doi.org/10.1007/s10530-015-0908-1

Naylor, R., Hindar, K., Fleming, I. A., Goldburg, R., Williams, S., Volpe, J., Whoriskey, F., Eagle, J., Kelso, D., & Mangel, M. (2005). Fugitive salmon: Assessing the risks of escaped fish from net-pen aquaculture. *BioScience, 55*(5), 427-437. https://doi.org/10.1641/0006-3568(2005)055[0427:FSATRO]2.0.CO;2

Naylor, R. L., Williams, S. L., & Strong, D. R. (2001). Aquaculture--A gateway for exotic species. *Science, 294*(5547), 1655-1656. https://doi.org/10.1126/science.1064875

Newman, W.A., & Ross, A. (1976). Revision of the balanomorph barnacles; including a catalog of the species. *Memoirs of the San Diego Society of Natural History, 9*, 1-108.

Novoa, A., Richardson, D. M., Pyšek, P., Meyerson, L. A., Bacher, S., Canavan, S., Catford, J. A., Čuda, J., Essl, F., Foxcroft, L. C., Genovesi, P., Hirsch, H., Hui, C., Jackson, M. C., Kueffer, C., Le Roux, J. J., Measey, J., Mohanty, N. P., Moodley, D., … & Wilson, J. R. U. (2020). Invasion syndromes: a systematic approach for predicting biological invasions and facilitating effective management. *Biological Invasions, 22*(5), 1801-1820. http://dx.doi.org/10.1007/s10530-020-02220-w

Obst, M., Exter, K., Allcock, A. L., Arvanitidis, C., Axberg, A., Bustamante, M., Cancio, I., Carreira-Flores, D., Chatzinikolaou, E., Chatzigeorgiou, G., Chrismas, N., Clark, M. S., Comtet, T., Dailianis, T., Davies, N., Deneudt, K., Diaz de Cerio, O., Fortič, A., Gerovasileiou, … & Pavloudi, C. (2020). A marine biodiversity observation network for genetic monitoring of hard-bottom communities (ARMS-MBON). *Frontiers in Marine Science, 7*. https://doi.org/10.3389/fmars.2020.572680

Occhipinti-Ambrogi, A. (2007). Global change and marine communities: alien species and climate change. *Marine Pollution Bulletin, 55*(7-9), 342-352. https://doi.org/10.1016/j.marpolbul.2006.11.014

Occhipinti-Ambrogi, A., & Galil, B. (2010). Marine alien species as an aspect of global change. *Advances in Oceanography and Limnology, 1*(1), 199-218. https://doi.org/10.1080/19475721003743876

Ojaveer, H., Galil, B. S., Carlton, J. T., Alleway, H., Goulletquer, P., Lehtiniemi, M., Marchini, A., Miller, W., Occhipinti-Ambrogi, A., Peharda, M., Ruiz, G. M., Williams, S. L., & Zaiko, A. (2018). Historical baselines in marine bioinvasions: Implications for policy and management. *PLoS One, 13*(8), e0202383. https://doi.org/10.1371/journal.pone.0202383

Oksanen., J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R.B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., Wagner, H. (2014). Package 'vegan'. https://cran.r-project.org/web/packages/vegan/vegan.pdf

Olenin S., Gollasch S., Lehtiniemi M., Sapota M., Zaiko A. (2017) *Biological invasions*. In: Snoeijs-Leijonmalm P., Schubert H., Radziejewska T. (eds). *Biological Oceanography of the Baltic Sea*. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-0668-2_5

OSPAR. (2019). Joint Harmonised Procedure for the Contracting Parties of HELCOM and OSPAR on the granting of exemptions under International Convention for the Control and Management of Ships' Ballast Water and Sediments, Regulation A-4. OSPAR Commission.

OSPAR. (2020). About OSPAR. OSPAR Commission. https://www.ospar.org/about

Peng, S., Kinlock, N. L., Gurevitch, J., & Peng, S. (2019). Correlation of native and exotic species richness: a global meta-analysis finds no invasion paradox across scales. *Ecology, 100*(1), e02552. https://doi.org/10.1002/ecy.2552

Radashevsky, V. I., Malyar, V. V., Pankova, V. V., Gambi, M. C., Giangrande, A., Keppel, E., Nygren, A., Al-Kandari, M., & Carlton, J. T. (2020). Disentangling invasions in the sea: molecular analysis of a global polychaete species complex (Annelida: Spionidae: Pseudopolydora paucibranchiata). *Biological Invasions, 22*(12), 3621-3644. https://doi.org/10.1007/s10530-020-02346-x

Raup, D. & R.E., Crick. (1979). Measurement of faunal similarity in paleontology. *Journal of Paleontology 53*(5), 1213-1227.

Read, G. B. (2010). Comparison and history of Polydora websteri and P. haswelli (Polychaeta: Spionidae) as mud-blister worms in New Zealand shellfish. *New Zealand Journal of Marine and Freshwater Research, 44*(2), 83-100. https://doi.org/10.1080/00288330.2010.482969

Rosenfield, D. C., & Coull, B. C. (1974). Adult morphology and larval development of Paramphiascella fulvofasciata n. sp. (Copepoda, Harpacticoida). *Cahiers de Biologie Marine, 15*(2), 295-317.

Rueckert, S., Wakeman, K. C., Jenke-Kodama, H., & Leander, B. S. (2015). Molecular

systematics of marine gregarine apicomplexans from Pacific tunicates, with descriptions of five novel species of Lankesteria. *International Journal of Systematic and Evolutionary Microbiology, 65*(8), 2598-2614. https://doi.org/10.1099/ijs.0.000300

Sagoff, M. (2018). What is invasion biology?. *Ecological Economics, 154*(C), 22-30. https://doi.org/10.1016/j.ecolecon.2018.07.023

Sardain, A., Sardain, E., & Leung, B. (2019). Global forecasts of shipping traffic and biological invasions to 2050. *Nature Sustainability, 2*(4), 274-282. https://doi.org/10.1038/s41893-019-0245-y

Sarno, D., Kooistra, W. H., Medlin, L. K., Percopo, I., & Zingone, A. (2005). Diversity in the genus Skeletonema (Bacillariophyceae). II. An assessment of the taxonomy of S. costatum-like species with the description of four new species. *Journal of Phycology, 41*(1), 151-176. https://doi.org/10.1111/j.1529-8817.2005.04067.x

Sato-Okoshi, W., Okoshi, K., Koh, B. S., Kim, Y. H., & Hong, J. S. (2012). Polydorid species (Polychaeta: Spionidae) associated with commercially important mollusk shells in Korean waters. *Aquaculture, 350*, 82-90. https://doi.org/10.1016/j.aquaculture.2012.04.013

Schlaepfer, M. A., Sax, D. F., & Olden, J. D. (2011). The potential conservation value of non-native species. *Conservation Biology, 25*(3), 428-437. https://doi.org/10.1111/j.1523-1739.2010.01646.x

Seebens, H., Gastner, M. T., Blasius, B., & Courchamp, F. (2013). The risk of marine bioinvasion caused by global shipping. *Ecology Letters, 16*(6), 782-790. https://doi.org/10.1111/ele.12111

Sharpe, A., Barrios, S., Gayer, S., Allan-Perkins, E., Stein, D., Appiah-Madson, H. J., Falco, R., & Distel, D. L. (2020). DESS deconstructed: Is EDTA solely responsible for protection of high molecular weight DNA in this common tissue preservative?. *PloS One, 15*(8), e0237356. https://doi.org/10.1371/journal.pone.0237356

Shaw, J. L., Clarke, L. J., Wedderburn, S. D., Barnes, T. C., Weyrich, L. S., & Cooper, A. (2016). 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

Shrader-Frechette, K. (2001). Non-indigenous species and ecological explanation. *Biology and Philosophy, 16*(4), 507-519. https://doi.org/10.1023/A:1011953713083

Slowikowski, K. (2021). *Ggrepel.* https://ggrepel.slowkow.com/index.html

Strayer, D. L. (2012). Eight questions about invasions and ecosystem functioning. *Ecology Letters, 15*(10), 1199-1210. https://doi.org/10.1111/j.1461-0248.2012.01817.x

Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology, 21*(8), 2045-2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x

Tait, L., Inglis, G., & Seaward, K. (2018). Enhancing passive sampling tools for detecting marine bioinvasions. *Marine Pollution Bulletin, 128*, 41-50. https://doi.org/10.1016/j.marpolbul.2018.01.015

Tang, C. Q., Leasi, F., Obertegger, U., Kieneke, A., Barraclough, T. G., & Fontaneto, D. (2012). The widely used small subunit 18S rDNA molecule greatly underestimates true diversity in biodiversity surveys of the meiofauna. *Proceedings of the National Academy of Sciences, 109*(40), 16208-16212. https://doi.org/10.1073/pnas.1209160109

Thomsen, M. S. (2020). Indiscriminate data aggregation in ecological meta-analysis underestimates impacts of invasive species. *Nature Ecology & Evolution, 4*(3), 312-314. https://doi.org/10.1038/s41559-020-1117-6

Tollington, S., Turbé, A., Rabitsch, W., Groombridge, J. J., Scalera, R., Essl, F., & Shwartz, A. (2017). Making the EU legislation on invasive species a conservation success. *Conservation Letters, 10*(1), 112-120. https://doi.org/10.1111/conl.12214

Trebitz, A. S., Hoffman, J. C., Darling, J. A., Pilgrim, E. M., Kelly, J. R., Brown, E. A., Chadderton, W. L., Egan, S. P., Grey, E. K., Hashsham, S. A., Klymus, K. E., Mahon, A. R., Ram, J. L., Schultz, M. T., Stepien, C. A., & Schardt, J. C. (2017). Early detection monitoring for aquatic non-indigenous species: Optimizing surveillance, incorporating advanced technologies, and identifying research needs. *Journal of Environmental Management, 202*, 299-310. https://doi.org/10.1016/j.jenvman.2017.07.045

Turbelin, A. J., Malamud, B. D., & Francis, R. A. (2017). Mapping the global state of invasive alien species: patterns of invasion and policy responses. *Global Ecology and Biogeography, 26*(1), 78-92. https://doi.org/10.1111/geb.12517

Valentini, A., Pompanon, F., & Taberlet, P. (2009). DNA barcoding for ecologists. *Trends in Ecology & Evolution, 24*(2), 110-117. https://doi.org/10.1016/j.tree.2008.09.011

Valéry, L., Fritz, H., Lefeuvre, J. C., & Simberloff, D. (2008). In search of a real definition of

the biological invasion phenomenon itself. *Biological Invasions, 10*(8), 1345-1351. https://doi.org/10.1007/s10530-007-9209-7

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

Verleye, T. J., De Raedemaecker, F., Vandepitte, L., Fockedey, N., Lescrauwaet, A.-K., De Pooter, D., & Mees, J. (Eds.). (2020). *Niet-inheemse soorten in het Belgisch deel van de Noordzee en aanpalende estuaria anno 2020*. Flanders Marine Institute. ISBN 9789464206005

Vitule, J. R., Occhi, T. V., Kang, B., Matsuzaki, S. I., Bezerra, L. A., Daga, V. S., Faria, L., de A. Frehse, F., Walter, F., & Padial, A. A. (2019). Intra-country introductions unraveling global hotspots of alien fish species. *Biodiversity and Conservation, 28*(11), 3037-3043. https://doi.org/10.1007/s10531-019-01815-7

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

Walday, M. & Kroglund, T. (2002). *The North Sea*. In: *Europe's biodiversity - biogeographical regions and seas (EEA Report No 1/2002)*. European Environmental Agency. https://www.eea.europa.eu/publications/report_2002_0524_154909/regional-seas-around-europe/NorthSea.pdf/view

Webber, B. L., & Scott, J. K. (2012). Rapid global change: implications for defining natives and aliens. *Global Ecology and Biogeography, 21*(3), 305-311. https://doi.org/10.1111/j.1466-8238.2011.00684.x

Weigle, S. M., Smith, L. D., Carlton, J. T., & Pederson, J. (2005). Assessing the risk of introducing exotic species via the live marine species trade. *Conservation Biology, 19*(1), 213-223. https://doi.org/10.1111/j.1523-1739.2005.00412.x

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

Williamson, M., & Fitter, A. (1996). The varying success of invaders. *Ecology, 77*(6), 1661-1666. https://doi.org/10.2307/2265769

Zaiko, A., Samuiloviene, A., Ardura, A., & Garcia-Vazquez, E. (2015). Metabarcoding approach for nonindigenous species surveillance in marine coastal waters. *Marine Pollution Bulletin, 100*(1), 53-59. https://doi.org/10.1016/j.marpolbul.2015.09.030

Zaiko, A., Schimanski, K., Pochon, X., Hopkins, G. A., Goldstien, S., Floerl, O., & Wood, S. A. (2016). Metabarcoding improves detection of eukaryotes from early biofouling communities: implications for pest monitoring and pathway management. *Biofouling, 32*(6), 671-684. https://doi.org/10.1080/08927014.2016.1186165

Addendum

Risk Assessment

Collecting the settlement plate and plankton samples posed minimal dangers to human health or the environment. Dimethyl sulfoxide is slightly irritating to the eyes, skin, and respiratory tract, while EDTA may cause serious irritation should it enter the eyes but is otherwise relatively harmless. Hence, specific precautions (such as gloves, eyewear, etc.) were not needed during collecting and morphological identification of the samples.

When settlement plate samples were homogenized, 10% bleach (sodium hypochlorite) solution was used to clean the equipment between processing of consecutive samples. Bleach solution poses several significant hazards: it releases toxic chlorine gas upon contact with acids, causes severe eye and skin burns, and poses an acute risk to aquatic organisms. Therefore, latex gloves were always worn to avoid skin contact, and the laboratory was kept well-ventilated. No specific waste treatment was needed, as only small quantities of bleach solution (applied to a tissue paper) were used at a time.

DNA extraction required the use of multiple chemical compounds, some of which posed health or environmental risks. Two of the solutions in the PowerSoil Pro kit (EA and C5) are classified as flammable by the manufacturer, and another two (CD1 and CD3) can form highly reactive compounds in the presence of bleach. However, neither hazard posed much risk as no bleach or spark-generating devices were present in the vicinity of these proprietary solutions during their use. CTAB (cetrimonium bromide) has the potential to irritate the skin and respiratory tract and to cause serious eye damage, and it is also highly toxic to the aquatic environment. Isopropanol and ethanol are both highly flammable. In addition, isopropanol has the potential to irritate the eyes, and its vapour can cause drowsiness if inhaled. Tris-EDTA buffer may irritate the eyes, respiratory tract, or skin if it comes into contact with them. However, the small volumes used (microliters to millilitres) and their transfer via a pipette meant that the above compounds posed little risk for contacting exposed body parts, especially since the usual laboratory precautions were followed at all times (wearing latex gloves and a lab coat and ensuring optimal ventilation of the room). In addition, waste fluids (such as the discarded supernatant from the DNA cleaning steps) were collected in a separate container for specialized chemical waste disposal. Sodium acetate and sodium chloride, also used during the DNA extraction, pose little risk to health or environment.

Chloroform, however, is significantly more hazardous than the other chemical compounds used for the DNA extraction. Not only does it irritate the skin, respiratory tract, and eyes upon contact, but it also is a suspected carcinogen and has the potential to damage organs if it enters the body. Despite its characteristic odour, detection of this odour occurs only above the

exposure limit to this chemical. For these reasons, all additions of chloroform to the DNA samples took place under a fume hood, and the waste from the chloroform-based DNA purification was collected separately for specialized chemical waste treatment.

For gel electrophoresis, DNA was visualized using Midori Green (NIPPON Genetics, Düren, Germany) and Gel Red (Merck, Darmstadt, Germany). These fluorescent dyes, due to their property of intercalating into DNA strands, were treated as potential mutagens: two pairs of latex gloves were worn at all times during the making and handling of electrophoretic gels, and all materials used in the preparation of these gels were kept in a dedicated corner of the laboratory and never moved outside this corner. Pipette tips and other single-use products that came into contact with the gel or the dyes were disposed of in a hazardous waste bin, as were the gels themselves. Other chemical compounds used in the preparation of the gels (loading dye containing bromophenol blue, tris-borate-EDTA buffer, agarose) posed minimal risk. Visualization of the bands on the gels required the use of ultraviolet light, which was applied to the gels in a closed chamber to avoid exposing ourselves to the potentially mutagenic wavelengths.

Table S1: Results of Nanodrop measurements of settlement plate community DNA

Sample	measurement	ng DNA /microliter	260/280	260/230
200817-UpUp	1st	16.6	1.63	0.27
200817-UpUn	1st	13.9	1.56	0.11
200817-LoUp	1st	20.7	1.80	0.18
200819-LoUn	1st	43.9	1.41	0.51
200819-LoUn	2nd	13.3	1.97	0.62
200819-LoUp	1st	10.0	1.88	0.14
200817-LoUn	1st	4.4	2.63	0.05
200819-UpUn	1st	13.7	1.86	0.09
200819-UpUp	1st	8.4	1.49	0.03
200820-UpUn	1st	17.7	1.89	0.79
200820-UpUp	1st	4.1	2.86	0.07
200817-LoUpA	1st	6.0	1.59	0.13
200817-LoUpB	1st	29.7	1.59	0.40
200819-UpUnC	1st	24.5	1.69	0.17
200819-UpUnD	1st	6.0	2.00	0.03
200819-LoUp	1st	4.4	1.42	0.12
200819-LoUp	2nd	3.9	1.97	0.13

	PO-UpUp	PO-UpUn	PO-LoUp	PO-LoUn	VD-UpUp	VD-UpUn	VD-LoUp	VD-LoUn	MM-UpUp
PO-UpUn	4.440699e-02								
PO-LoUp	3.556591e-03	5.862543e-01							
PO-LoUn	6.666947e-04	8.989621e-02	1.395291e-01						
VD-UpUp	9.042528e-01	8.197136e-01	8.650429e-01	9.704007e-01					
VD-UpUn	1.681587e-01	4.102760e-02	9.221373e-01	1.399003e-01	7.460116e-01				
VD-LoUp	1.684059e-01	8.989621e-02	8.076402e-01	5.101538e-01	2.235166e-01	1.007690e-03			
VD-LoUn	6.882263e-01	5.912479e-01	7.873031e-01	5.912479e-01	4.470355e-01	6.678975e-05	4.102760e-02		
MM-UpUp	1.681587e-01	1.399003e-01	5.710065e-01	1.399003e-01	7.460116e-01	6.219246e-03	8.070027e-03	1.091607e-01	
MM-UpUn	1.273321e-01	2.454882e-02	4.508417e-01	9.595690e-02	6.961823e-01	1.599825e-02	2.454882e-02	6.594100e-02	2.746430e-04

Figure S2: Distance matrix of the morphological settlement plate data

	PO-UpUp	PO-UpUn	PO-LoUp	PO-LoUpA	PO-LoUn	VD-UpUp	VD-UpUn	VD-LoUp	VD-LoUn	MM-UpUp
PO-UpUn	0.76078202									
PO-LoUp	0.76765974	0.28755251								
PO-LoUpA	0.74420888	0.21794198	0.10340065							
PO-LoUn	0.80212354	0.36208390	0.12278576	0.21356855						
VD-UpUp	1.00000000	0.94117568	0.99938402	1.00000000	0.99883183					
VD-UpUn	0.99815001	0.97695372	0.99933218	1.00000000	0.99863636	0.98019069				
VD-LoUp	0.99833010	0.94254284	0.99942025	1.00000000	0.99884637	0.15282715	0.98134951			
VD-LoUn	1.00000000	0.93201813	1.00000000	1.00000000	0.99883503	0.88307148	0.20191663	0.88532193		
MM-UpUp	0.99880108	0.92846695	0.99930245	1.00000000	0.99771835	0.48717005	0.90685431	0.49680900	0.80857624	
MM-UpUn	1.00000000	0.93716844	1.00000000	1.00000000	0.99797573	0.10154986	0.97735399	0.08440971	0.87960702	0.44424495

Figure S3: Distance matrix of the 18S rRNA settlement plate data

	PO-6-100	PO-6-300	MM-6-100	MM-6-300	VD-6-100	VD-6-300	PO-8-100	PO-8-300	MM-8-100	MM-8-300	VD-8-100
PO-6-300	4.225535										
MM-6-100	10.291998	8.708301									
MM-6-300	6.899879	4.202616	7.868905								
VD-6-100	13.290530	12.105680	13.338297	11.077828							
VD-6-300	10.713688	9.250033	11.301626	8.235033	9.338916						
PO-8-100	8.467997	6.451067	9.105830	5.263214	12.394721	9.669980					
PO-8-300	9.397345	7.630162	9.975934	6.656166	13.047307	10.493409	7.865398				
MM-8-100	8.799926	6.880987	9.415299	5.782140	12.623820	9.961939	7.118696	7.099636			
MM-8-300	8.793367	6.872597	9.409169	5.239985	11.864324	9.149663	6.829067	6.004889	6.538001		
VD-8-100	8.027205	5.860487	8.697424	4.520008	12.097858	9.286418	6.354541	7.371646	6.384446	6.625913	
VD-8-300	7.644928	5.324867	8.345902	3.799917	11.847661	8.958039	5.420924	6.749357	6.062103	5.961741	2.011747

Figure S4: Distance matrix of 18S rRNA plankton data

	PO-6-100	PO-6-300	MM-6-100	VD-6-100	VD-6-300	PO-8-100	PO-8-300	MM-8-100	MM-8-300	VD-8-100
PO-6-300	6.435929									
MM-6-100	9.001661	8.461235								
VD-6-100	8.844584	8.293930	9.322603							
VD-6-300	7.955473	7.338403	8.852869	5.085283						
PO-8-100	6.803923	6.070911	7.834354	7.653357	6.605813					
PO-8-300	5.743894	4.853341	6.933690	6.710055	5.491052	3.609314				
MM-8-100	7.900999	7.279313	8.803950	8.590542	7.687300	5.375069	5.373768			
MM-8-300	7.212879	6.525965	8.192030	7.744805	6.794078	5.134870	4.285929	5.392354		
VD-8-100	7.970348	7.354526	8.245149	8.127960	7.736912	6.043104	5.518930	7.611908	6.900556	
VD-8-300	6.559230	5.795352	7.622811	6.898629	5.885752	4.834327	3.069750	6.093806	4.223004	6.087476

Figure S5: Distance matrix of COI plankton data

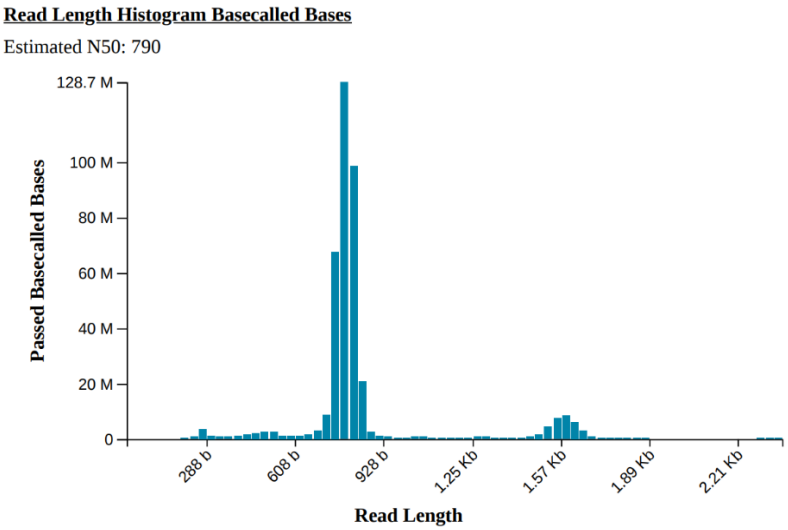

Figure S6: Read length distribution of the first Nanopore run

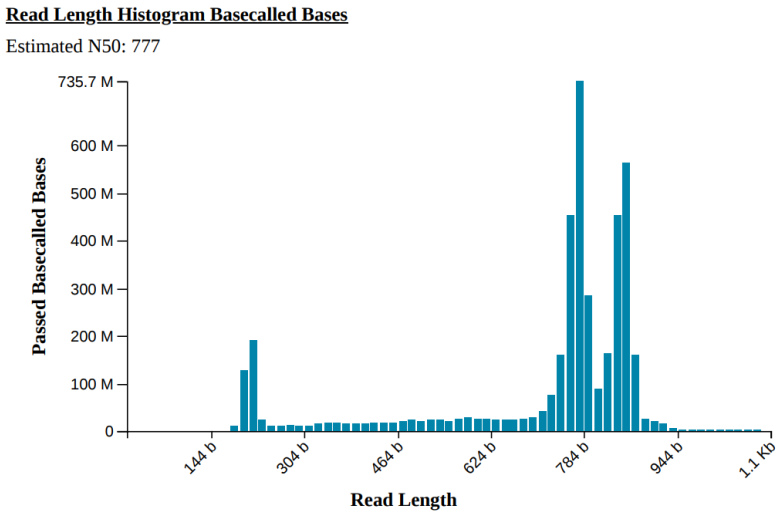

Figure S7: Read length distribution of the second Nanopore run

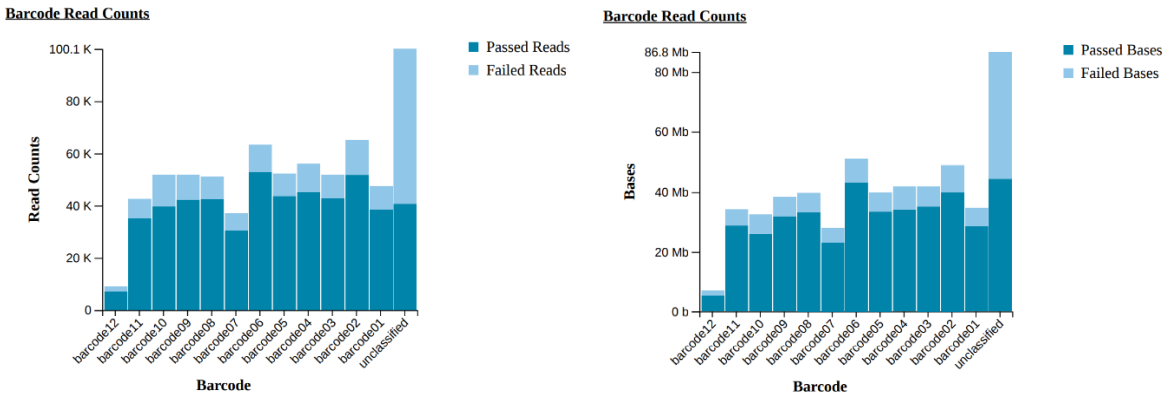

Figure S8: Proportion of passed reads and bases of the first Nanopore run

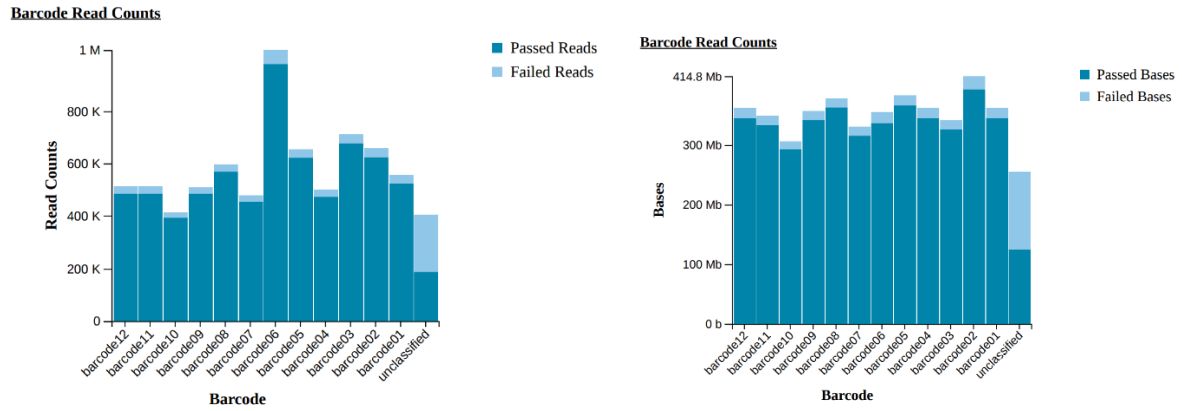

Figure S9: Proportion of passed reads and bases of the second Nanopore run

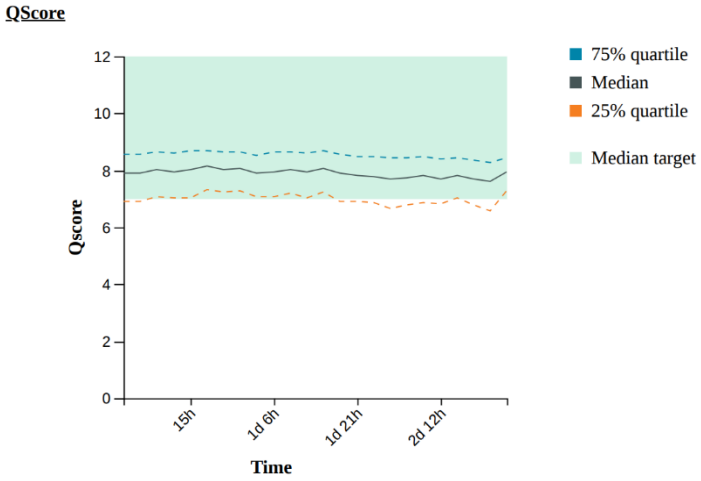

Figure S10: Quality score in function of time for the first Nanopore run

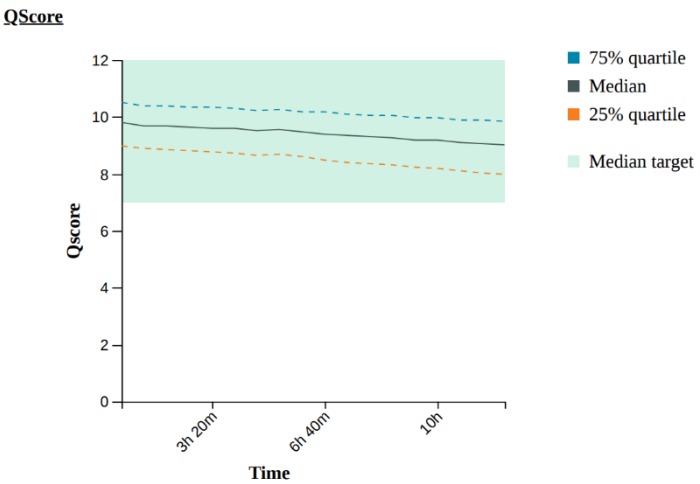

Figure S11: Quality score in function of time for the second Nanopore run

AFDELING ECOLOGIE, EVOLUTIE EN BIODIVERSITEITSBEHOUD
Charles Deberiotstraat 32 bus 2439
3000 LEUVEN, BELGIË
tel. + 32 16 32 39 66
fax + 32 16 32 45 75
www.bio.kuleuven.be/eeb