Biol Invasions (2025) 27:77
https://doi.org/10.1007/s10530-024-03503-2

ORIGINAL PAPER

Using the long-term genetic monitoring network ARMS-MBON to detect marine non-indigenous species along the European coasts

Justine Pagnier · Nauras Daraghmeh · Matthias Obst

Received: 25 July 2024 / Accepted: 26 November 2024 / Published online: 1 February 2025

Abstract The increasing prevalence of non-indigenous species (NIS) in marine ecosystems poses significant challenges for biodiversity conservation and ecosystem management. Advances in molecular techniques enable early detection and long-term monitoring of such taxa, especially when coupled with spatio-temporally wide sampling by networks such as the European ARMS Marine Biodiversity Observation Network (ARMS-MBON). This initiative performs standardised sampling campaigns using autonomous reef monitoring structures (ARMS) along European coasts and adjacent regions, providing open-access DNA metabarcoding data sets. We tested the potential of genetic observatory networks to detect and monitor marine NIS by analysing all publicly available ARMS-MBON cytochrome c oxidase subunit I (COI) and 18S rRNA amplicon sequencing data as of February 2024 using a customised bioinformatic pipeline. Screening against the World Register of Introduced Marine Species (WRiMS) and applying manual curation, we identified 63 marine taxa considered non-indigenous at one or more locations. This included widespread taxa and potential new introductions, such as *Eucheilota menoni* in the Adriatic Sea. We found no significantly higher number of NIS in samples from locations particularly impacted by maritime traffic compared to other areas. Our results suggest that the genetic observatory network approach is powerful for detecting and monitoring marine NIS, and that manual curation still is an essential step for obtaining reliable results. We recommend key improvements including more spatially intense sampling across diverse environments as well as enhancement of NIS reference checklists and genetic databases to ensure accurate identification of both known and unknown NIS across Europe.

Keywords Alien species · Invasive species · Autonomous reef monitoring structures (ARMS) · COI · 18S rRNA · Amplicon sequencing

Justine Pagnier and Nauras Daraghmeh authors have contributed equally.

Justine Pagnier: The research was initiated and conducted while the author was employed by LifeWatch ERIC, Sector II-III, Plaza de España, 41071, Seville, Spain

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10530-024-03503-2.

J. Pagnier (✉) · M. Obst
Department of Marine Sciences, University of Gothenburg, Box 461, 405 30 Gothenburg, Sweden
e-mail: justine.pagnier@gu.se

J. Pagnier · N. Daraghmeh · M. Obst
Gothenburg Global Biodiversity Centre (GGBC), University of Gothenburg, Box 463, 405 30 Gothenburg, Sweden

N. Daraghmeh
Department of Biological and Environmental Sciences, University of Gothenburg, Box 463, 40 530 Gothenburg, Sweden

Introduction

Marine ecosystems play crucial environmental and socio-economic roles, yet they face significant threats such as biodiversity loss and degradation (Luypaert et al. 2020; Mazaris et al. 2019; O'hara et al. 2021), often exacerbated by non-indigenous species (NIS). These species, usually introduced through global maritime traffic, pose significant risks to native marine communities (De Poorter et al. 2008; Katsanevakis et al. 2023; Pyšek et al. 2020; Tsiamis et al. 2020). Rapid identification and management of NIS are crucial, especially when introduced populations are still localised and manageable (Pyšek et al. 2020; Trebitz et al. 2017).

In the scientific literature, various terms are used—often interchangeably—to refer to introduced taxa in marine systems, oftentimes describing different stages of establishment or levels of impact. These terms include invasive species, alien species, introduced species, non-native species, or neobiota (Kuhlenkamp and Kind 2018). We here use the term NIS throughout this study to refer to taxa introduced to marine regions outside of their known native range likely through anthropogenic activities. This term does not necessarily imply a potential harmful status of a taxon (often termed invasive in this context), which may yet to be determined for various NIS.

Historically, marine NIS have been recorded through visual surveys based on morphological identifications, which requires taxonomic expertise and tends to overlook small and cryptic organisms (Leray and Knowlton 2016). Recently, DNA-based methods have emerged as an innovative tool for the detection and identification of NIS (Ammon et al. 2018; Coster et al. 2021; Darling and Mahon 2011; Fonseca et al. 2023; Holman et al. 2019; Othman et al. 2023; Pearman et al. 2020; Wu et al. 2023). Combining DNA metabarcoding with high-throughput sequencing (HTS) represents a non-disruptive approach utilising unique genetic signatures present in environmental or bulk community samples to determine taxonomic composition of biotic communities and screen rapidly for the presence of potentially harmful organisms (Hablützel et al. 2023). While in this context metabarcoding often refers to the analysis of environmental DNA (eDNA) present in soil, water, or air, it can also be applied to other genetic material, such as whole organism community DNA (wocDNA) (Creedy et al. 2022) or DNA from gut contents (Trujillo-González et al. 2022) and faecal samples (McInnes et al. 2017). Molecular techniques also tackle the challenges posed by morphological identification of species at early life stages, such as the egg and larval stage (Gallage et al. 2023; Hoffman et al. 2021), and through their ability to detect rare and elusive species (Goldberg et al. 2016). In fact, molecular methods are extremely sensitive, able to detect DNA of NIS and other rare species making up only a small fraction of total biomass in a sample (Hatzenbuhler et al. 2017).

To this date, despite the immense potential of DNA metabarcoding for marine invasion ecology, studies applying this method across large spatial scales remain relatively scarce. Previous works have analysed eDNA to assess marine metazoan diversity in a regional context (Fraija-Fernández et al. 2020; West et al. 2021). Recently, the *Tara* Oceans initiative—through its global sampling campaigns and provision of resulting metabarcoding data sets—has enabled groundbreaking research on marine plankton communities on a planetary scale (Bork et al. 2015). Brown et al. (2016) were the first to apply metabarcoding on marine and freshwater samples to screen for NIS on a (semi)continental scale but this approach has not yet been integrated into long term ecological research. However, in the context of global change and resulting species migrations, large-scale studies of marine biodiversity and invasion patterns are essential to understand global ecosystem dynamics (Lacoursière-Roussel et al. 2018).

Previous research on NIS in aquatic environments typically employed molecular techniques aimed at detecting specific species like the American bullfrog (*Lithobates catesbeianus*) (Dejean et al. 2012), Asian carps (*Hypophthalmichthys nobilis* and *Hypophthalmichthys molitrix*) (Jerde et al. 2013; Mahon et al. 2013), the bluegill sunfish (*Lepomis macrochirus*) (Takahara et al. 2013), the New Zealand mudsnail (*Potamopyrgus antipodarum*) (Goldberg et al. 2013), and the red swamp crayfish (*Procambarus clarkii*) (Tréguier et al. 2014). Although these methods are efficient for tracking well-known invaders, they fail to account for other potentially significant NIS (Lawson Handley 2015). A more beneficial approach to NIS management could involve passive surveillance based on the description of entire communities through metabarcoding.

A number of studies have used metabarcoding to detect NIS among whole marine communities, however they have primarily focused on limited geographic areas (Borrell et al. 2017; Günther et al. 2018; Westfall et al. 2020; Zarcero et al. 2024) or used customised reference libraries tailored to known NIS (Couton et al. 2019), thereby limiting the detection to already known species in an area. In such spatially and taxonomically more limited contexts, it is common to compile a preliminary list of target taxa to screen for, which can be more challenging for studies on a larger geographical scale. Previous studies have relied on such databases to screen for NIS on a national scale (Holman et al. 2019), while Grey et al. (2018) for example, used previous records from databases such as the World Register of Marine Species (WoRMS) (Ahyong et al. 2024) and the IUCN Red List to assess if a putative NIS truly was introduced in a specific area.

Developing global and curated databases that catalogue NIS and their known introduction regions is crucial for large-scale screening efforts. Several initiatives, among others, have been contributing to this goal. These include the World Register of Introduced Marine Species (WRiMS; Costello et al. 2021), which is based on the larger and widely recognized WoRMS; the marInvaders web application (Verones et al. 2023), which integrates data from the Ocean Biodiversity Information System (OBIS; Intergovernmental Oceanographic Commission of UNESCO 2024), WoRMS, the Global Invasive Species Database (GISD; Invasive Species Specialist Group ISSG 2015), and a database from the Nature Conservancy (Molnar et al. 2008); the Global Register of Introduced and Invasive Species (GRIIS; Pagad et al. 2018); and the AquaNIS database (AquaNIS, 2015).

Various methodologies have been used to sample DNA in terrestrial and aquatic environments (Thomsen and Willerslev 2015). Autonomous reef monitoring structures (ARMS) are standardised, easy-to-produce devices used for non-destructive sampling of marine benthic communities. These structures were originally designed to mimic natural coral reef habitats to facilitate the collection of genetic material from cryptobenthic reef communities (Knowlton et al. 2010; Leray and Knowlton 2015; Plaisance et al. 2011) but are now applied in a variety of marine ecosystems (Obst et al. 2020; Thomasdotter et al. 2023; Cecchetto et al. 2024). These sampling units offer a unique opportunity to study marine biodiversity on hard-bottom substrates and to detect NIS in a standardised and replicable manner by obtaining bulk samples of the communities that inhabit these structures.

In this study, we aim to explore the potential of a standardised sampling approach with DNA metabarcoding for detecting marine NIS on hard-bottom substrates and studying their potential range expansion or new introduction across broad geographic scales using amplicon sequencing raw data generated by the European ARMS Marine Biodiversity Observation Network (ARMS-MBON) (Obst et al. 2020). We hypothesise that genetic observatory networks, as implemented in ARMS-MBON, have the potential to detect specific non-indigenous species across diverse geographic locations and time points. This network continuously deploys ARMS in the vicinity of marine stations, ports, marinas, marine protected areas (MPAs), and long-term ecological research (LTER) sites distributed along the shores of Europe as well as in additional subtropical and polar regions. The raw and processed genetic data are then made open access to the public, as well as image data and important metadata (Daraghmeh et al. 2024). We first harvested all publicly available genetic data (as of February 2024) of two universal marker genes from this network, which encompasses the entirety of ARMS-MBON sampling campaigns conducted from its initiation in 2018 up to the year 2021. We screened the processed data against the WRiMS database, because of its comprehensiveness and ability to filter the precise regions of the observatories (see detailed methods below). Furthermore, we hypothesise that manual curation of the resulting species lists is essential to ensure accuracy when using such databases, not only because of potential limitations in automatic taxonomic assignment and the occurrence of erroneous detections but also due to the inherent limitations of NIS databases themselves.

To the best of our knowledge, our study represents (i) the largest set of ARMS metabarcoding samples analysed to date; (ii) the first work using ARMS MBON metabarcoding data solely to comprehensively screen for NIS; and (iii) a NIS scan across one of the largest geographical areas to date. By analysing data of two universal molecular markers (i.e., one mitochondrial and one nuclear gene) and developing a custom pipeline for efficient data processing and manual taxonomic

occurrence curation, we investigate on the potential of the genetic observatory network approach for the early detection and continuous monitoring of marine NIS. This study does not aim to provide a quantitative assessment of detection efficacy but instead offers a case study evaluating the capability of the genetic observatory approach for broad-scale NIS detection. We further describe the limitations encountered, suggest points for improvement, and give recommendations for future applications.

Materials and methods

Spatial and temporal coverage

The data set used for this study was generated by the European ARMS-MBON programme (Obst et al. 2020). It encompasses samples collected by 19 observatories distributed across 14 countries, which deployed (and continuously deploy) ARMS units at one to seven sites each, in the vicinity of marinas,

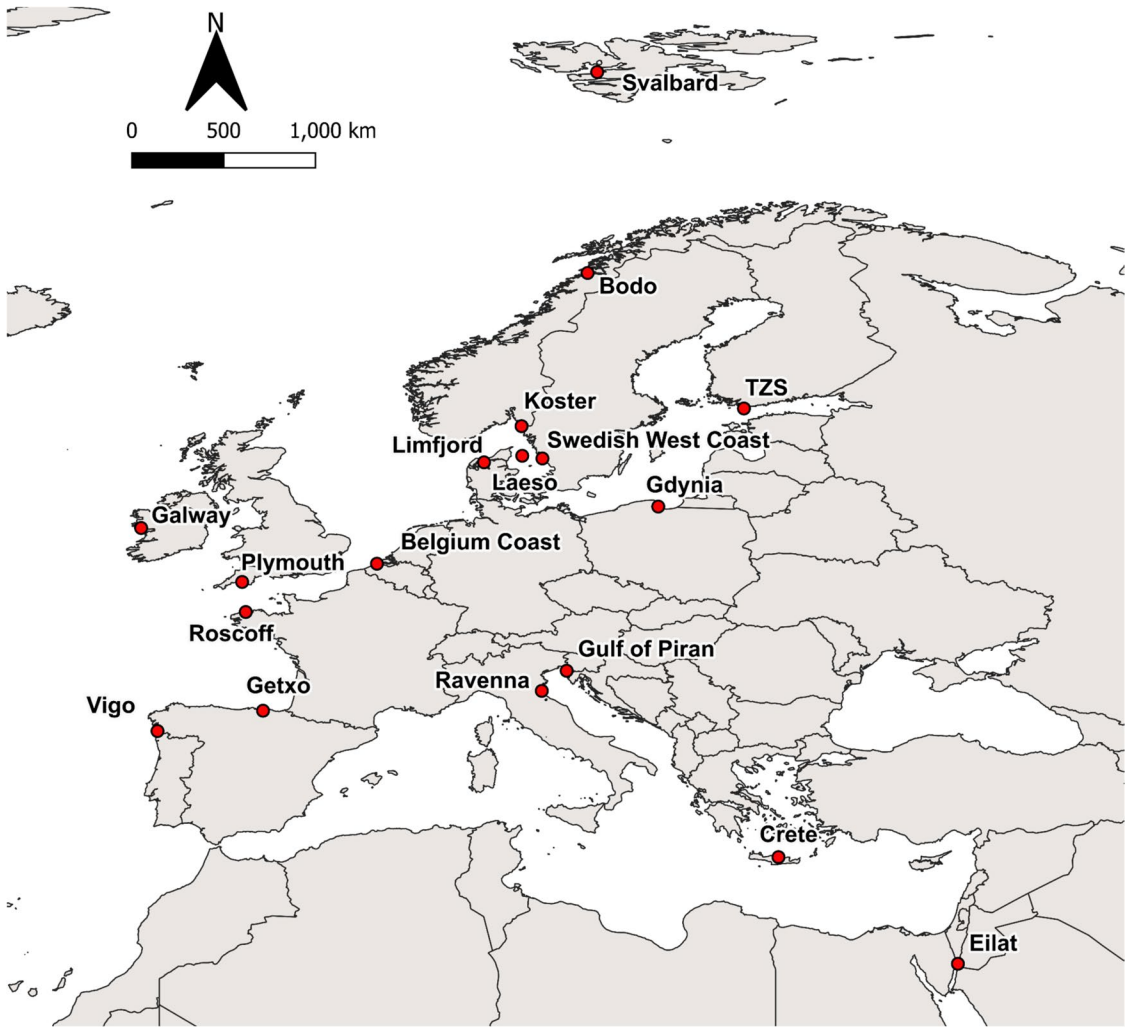

Fig. 1 Location of the 19 ARMS-MBON observatories which collected samples used in this study. The two Ravenna (Italy) observatories are presented as a single entity here due to their proximity. TZS—Tvärminne Zoological Station. Map produced using *QGIS* v3.34

ports, MPAs, and LTER sites—from Svalbard in the north to the Red Sea in the south (Fig. 1). This comprises 112 ARMS units in total, which remained deployed for varying periods between April 2018 and November 2021. The duration of sampling events (i.e., deployment duration of individual ARMS units) ranges from 37 to 649 days. Observatories processed all ARMS units and for each unit preserved homogenised samples of the following three fractions: i) one sessile fraction (all 40-µm-filtered sessile material attached to ARMS plates, hereafter referred to as "sessile fraction"); and ii) two motile fractions representing all biological material not attached to ARMS plates, i.e., one fraction of material with a size of 100 to 500 µm (hereafter "motile fraction 100 µm") and one fraction of material larger than 500 µm (hereafter "motile fraction 500 µm"). Details on observatories and deployments can be found in Supplementary File 1. The descriptions of observatory design and fieldwork procedures have been expanded upon since Obst et al. (2020). ARMS units were deployed, retrieved, photographed, and processed by individual partners in the ARMS-MBON network following the ARMS-MBON protocols (accessible through the ARMS-MBON GitHub repository at https://github.com/arms-mbon). See Table 1 for a list of the institutions and investigators who performed the sampling and Supplementary Table S1 for links to ARMS-MBON webpages and GitHub repositories.

Description of the amplicon sequencing data set

Molecular laboratory procedures (i.e., DNA extraction, PCR amplification and amplicon sequencing) for samples from ARMS units processed by each network partner (see above) were carried out centrally at the facilities of an ARMS-MBON network partner, the Hellenic Centre for Marine Research (HCMR), Crete, Greece. Protocols and documentation on this are accessible via the ARMS-MBON GitHub repository (https://github.com/arms-mbon), see Supplementary Table S1 for links. We obtained all publicly available (as of February 17th, 2024) amplicon sequencing raw data of the mitochondrial cytochrome *c* oxidase subunit I (COI) and nuclear 18S rRNA (18S) marker genes from ARMS-MBON. These markers amplify the 313-bp-long "Leray fragment" of COI (Geller et al. 2013; Leray et al. 2013) and a fragment of the hypervariable V9 region of 18S (Hardy et al. 2010). Contrary to the descriptions by Hardy et al. 2010, we observed that most amplicons of the ARMS-MBON 18S data set were shorter than 200 bp.

In total, the data set encompasses 349 sequencing samples for COI (plus nine negative control samples) and 354 sequencing samples for 18S (plus eight negative control samples). Raw sequences of all samples are available on the European Nucleotide Archive (ENA; https://www.ebi.ac.uk/ena/browser/home) (Yuan et al. 2024) and accession numbers can be found in Supplementary File 2. We refer to the computational protocol described below (Daraghmeh 2024) for details on how raw data was obtained for this study.

Procedure for NIS detection

A computational protocol with detailed descriptions of bioinformatics processing of sequencing data and subsequent NIS identification has been published (Daraghmeh 2024). We briefly describe the procedure below and refer to the published protocol for all code and Supplementary Text S1 for details on software used and parameter settings applied.

Sequence denoising, clustering into molecular operational taxonomic units and taxonomic classification

We processed COI and 18S raw reads separately in *R* v4.3.1 (R Core Team 2021). We used *cutadapt* v4.5 (Martin 2011) to remove primers from sequences still containing them (some of the demultiplexed sequence reads provided by ARMS-MBON were already devoid of primers). Quality assessment and read trimming were conducted using *DADA2* v1.28.0 (Callahan et al. 2016) prior to constructing error models and inferring amplicon sequence variants (ASVs). We subsequently merged paired-end reads and removed chimeric sequences via *DADA2*'s *removeBimeraDenovo* function. Read numbers remaining in each sample were tracked through each step of the pipeline (see Supplementary File 3).

Taxonomic classification for 18S ASVs was performed within *DADA2* via the *RDP* classifier (Wang et al. 2007), using the *SILVA* v132 reference set (Quast et al. 2013) formatted for use in *DADA2* (Callahan 2018) and a subset of its eukaryotic sequences clustered at 99% similarity (Morien and Parfrey

Table 1 Contributing institutions and individuals involved in ARMS deployment, retrieval and processing for each observatory

Observatory	Contributing institutes	Investigators
Belgium Coast	Flanders Marine Institute, VLIZ, Oostende, Belgium	Rune Lagaisse, Klaas Deneudt, Sander Delacauw, Jonas Mortelmans
Bodo	Nord University, Bodo, Norway	Henning Reiss
Crete	Hellenic Centre for Marine Research, Crete, Greece	Thanos Dailianis, Vasilis Gerovasileiou, Christina Pavloudi, Eva Chatzinikolaou, Markos Digenis, Jon Bent Kristoffersen, Emmanouela Vernadou, Giorgos Chatzigeorgiou
Galway	Plentzia Marine Station (PiE-UPV/EHU), University of the Basque Country, Spain	Anne Marie Power, Louise Allcock
Gdynia	Institute of Oceanology Polish Academy of Sciences, University of Gdansk, Poland	Rafal Lasota, Maciej Chetchowski
Getxo	Plentzia Marine Station (PiE-UPV/EHU), University of the Basque Country, Spain	Ibon Cancio, Oihane Diaz de Cerio, Javier Tajadura
Gulf of Piran	Marine Biology Station Piran, National Institute of Biology (NIB), Slovenia	Borut Mavric
Eilat	Interuniversity Institute for Marine Sciences, Eilat, Israel	Liraz Levy, Amatzia Genin
Koster	Department of Marine Sciences, University of Gothenburg, Sweden	Matthias Obst
Laeso Limfjord	Department of Ecoscience, Aarhus University, Denmark	Peter A.U. Staehr
Plymouth	Marine Biological Association of the UK, The Laboratory, Citadel Hill, Plymouth, UK	Nathan Chrismas
Ravenna Harbour	Department of Biological, Geological and Environmental Science, University of Bologna, Bologna, Italy	Federica Costantini, Alessandro Piazza
Ravenna Marina	Department of Biological, Geological and Environmental Science, University of Bologna, Bologna, Italy	Federica Costantini, Alessandro Piazza
Roscoff	Station Biologique De Roscoff, Sorbonne University, France	Frederique Viard, Thierry Comtet
Svalbard	Institute of Oceanology Polish Academy of Sciences, University of Gdansk, Poland	Matgorzata Zbawicka, Magdalena Matachowicz, Anita Pocwierz-Kotus, Piotr Kuklihski, Piotr Balazy
Swedish West Coast	Department of Marine Sciences, University of Gothenburg, Sweden	Matthias Obst
TZS	Tvarminne Zoological Station, Faculty of Biological and Environmental Sciences, University of Helsinki, Finland	Laura Kauppi, Jostein Solbakken
Vigo	Centra de Investigation Marina, Universidade de Vigo, Spain	Jose Gonzalez, Estefania Paredes, Jesus Souza Troncoso

2018); as well as the *PR2* v5.0.0 (Guillou et al. 2013) database. An ensemble taxonomy was generated for 18S ASVs using parts of the *ensembleTax* v1.2.2 (Catlett et al. 2023) package, merging the taxonomy tables of the three different reference assignments.

Putative nuclear mitochondrial pseudogenes (NUMTs) were identified using *MACSE* v2.05 (Ranwez et al. 2018) in *R* v4.1.0 and removed from the COI data set. We applied a negative-control correction to remove all ASVs for which the read count within negative control samples amounted to 10% or more of their total read count. Subsequently, clustering of ASVs into molecular operational taxonomic units (mOTUs) was performed using *swarm* v3.0.0 (Mahé et al. 2015), setting the parameter *d* to 13 and 1 for COI and 18S, respectively. As described by

Mahé et al. (2014), the use of high *d*-values may be considered when working with relatively fast evolving marker genes such as COI. A comparatively high *d*-value of 13 has been identified as most suitable for the "Leray fragment" (Antich et al. 2021) and has now been used widely in studies applying COI amplicon sequencing (Antich et al. 2021; Atienza et al. 2020; Bakker et al. 2019; Siegenthaler et al. 2019; Wangensteen et al. 2018). The resulting mOTUs were further curated using the *R* package *LULU* v0.1.0 (Frøslev et al. 2017) to remove putative erroneous mOTUs after generating a match list using *BLAST* + v2.11.0 (Camacho et al. 2009).

Taxonomy was assigned to COI mOTUs using two reference databases via two classification tools. We queried against the Barcode of Life Data System (BOLD) (Ratnasingham and Hebert 2007) database using *BOLDigger-commandline* v2.2.1 (Buchner and Leese 2020). In addition, we assigned taxonomy to our COI data set using MIDORI2 (Leray et al. 2022) via the associated web server tool *MIDORI server* (Leray et al. 2018). The latter was based on *GenBank* release 257 (Benson et al. 2012) at the time of access (February 2024). In *R*, we then compared classification based on *MIDORI2* and *BOLD* and created a final taxonomy table using packages from the *tidyverse* v1.3.1 collection (Wickham et al. 2019). Ultimately, both COI and 18S mOTUs with the same species classification were merged and the corresponding read abundances summed. Additional *R* packages used during all previously described steps include *ggplot2* v3.4.2 and v3.4.3 (Wickham 2016), *Biostrings* v2.68.1 (Pagès et al. 2020), *ShortRead* v1.58.0 (Morgan et al. 2009), *hiReadsProcessor* v1.29.1 and v1.36.0 (Malani 2021), *dplyr* v1.0.9 and v1.1.3 (Wickham, François, et al., 2023), *tidyr* v1.3.0 (Wickham, et al., 2023), *seqinr* v4.2.30 (Charif and Lobry 2007), *argparse* v2.2.2 (Davis 2023), *stringr* v1.5.0 (Wickham 2022), *devtools* v2.4.3 (Wickham et al. 2022), and *remotes* v2.4.2 (Csárdi et al. 2024).

Screening for non-indigenous species

To screen for NIS across our data set, we chose to base our analyses on the WRiMS database because of its focus on marine non-native species, and it is part of the larger and globally recognized World Register of Marine Species (WoRMS).

We queried against the WRiMS database (Costello et al. 2024) to obtain a list of putative NIS to screen for. Here, we downloaded lists of species registered in WRiMS for the regions corresponding to the observatory locations (accessed on February 26th, 2024). For details on how the standardised search was performed, see Daraghmeh, (2024). All individual lists were then merged in *R* using packages mentioned above and the *xlsx* v0.6.5 package (Dragulescu and Arendt 2020) to obtain a list of taxa registered in WRiMS for regions corresponding to the ARMS-MBON sampling locations.

We then filtered the COI and 18S data sets in *R* using the packages *phyloseq* v1.36.0 (McMurdie and Holmes 2013) and *data.table* v1.14.2 (Dowle and Srinivasan 2021) to extract mOTUs classified at the species level. Species-level assignments were matched against the World Register of Marine Species (WoRMS) database (Ahyong et al. 2024) to obtain the respective accepted names for each species using the web services of LifeWatch Belgium (https://www.lifewatch.be/data-services/). All mOTUs assigned to a species found in the compiled WRiMS list were subset with their respective occurrence information. We then tracked all ASVs throughout the pipeline to obtain information on which of these putative NIS mOTUs they ended up in. This was done to manually check every NIS occurrence (see below). Occurrences of ASVs with less than 5 reads were set to zero. We also briefly assessed which additional NIS would be detected if no read threshold were to be applied (however, the data not filtered based on read number were not further considered in downstream analysis). All size fraction samples for each ARMS sampling event (i.e., an individual ARMS unit deployed for a given period) were merged and corresponding read abundances were summed.

Manual curation of NIS occurrences

Taxonomic classification of sequences from DNA metabarcoding remains challenging, as the markers applied may lack sufficient species-level resolution and/or the amplicons may be too short to obtain species-level classification with a high confidence (Porter and Hajibabaei 2020). We meticulously curated potential NIS occurrences. For the COI marker, each ASV found in a putative NIS mOTU was again manually queried against BOLD (or NCBI's GenBank in

some cases). For 18S, ASVs were manually queried against GenBank (or PR2 in some cases). Additionally, we double-checked all occurrences to ensure that only species genuinely considered non-indigenous and potentially present in the studied regions were included. This extra curation step was necessary because we relied on WRiMS, which, while valuable, might have erroneous entries for some regions. As a result, we had to further validate the species to ensure accurate identification and avoid false positives, ensuring that the listed alien species were both biologically plausible for the region and aligned with reliable references.

During this process, we applied strict criteria for ASV inclusion:

1. For COI, we discarded ASVs if (i) a species level match could not be made according to BOLD and one of the likely species assignments given by BOLD was a native species; or (ii) there was no match in BOLD and where for these ASVs classification on GenBank had no hits with a query cover of 100% and similarity above 98%.
2. For 18S, ASVs were removed if (i) the highest valued match (i.e., based on E-value) in NCBI had no hits with a query cover of 100% and similarity above 98% for the mOTU's species assignment and there was no hit in PR2 with similarity above 98%; or (ii) the top hits with the highest E-value in NCBI with a query cover of 100% and similarity above 98% represented several taxa and based on the ASV's occurrence and the scientific literature it could not be determined if the sequence likely belonged to the respective NIS.
3. For both genes, occurrences were set to zero if (i) all occurrences of the respective taxon appeared at locations where it can be considered native (this means the curated data set still contained occurrences of mOTUs where they are considered native, as long as they occurred at least at one location where they are a putative NIS); or (ii) there was no clear indication based on scientific literature that a taxon was potentially considered as a NIS at any of the locations of occurrence; or (iii) if the literature review provided evidence that a taxon was highly unlikely to be present in a given area, e.g., taxa with little tolerance to low salinity found the Baltic Sea. The latter case occurred in five taxa and warrants further study.

Ultimately, curated data for COI and 18S markers were combined to generate a comprehensive data set detailing NIS distribution across sampling events. Details of this curation process, including thresholds and exclusion criteria, are provided in Supplementary Text S1 and Supplementary File 5.

Comparison to other NIS databases

The NIS list we obtained from WRiMS for our study area (see above) was compared to NIS lists from other databases for the same, or as similar as possible, area. We built 3 other lists (see Supplementary Table S3 for additional information on how they were built):

- with the AquaNIS database (407 species);
- with the marInvaders database (450 species);
- with the GRIIS database (943 species).

All species lists were obtained for comparable marine regions, except for the GRIIS database, which classifies NIS at the country level. To visualise the overlaps and differences between these NIS lists, we created an UpSet plot (Supplementary Fig. S11). Next, we examined how many of the species we detected (COI and 18S combined) were present in each database and how the databases complemented one another (Supplementary Fig. S12).

Analysis of NIS observations across the network

We initially assessed a data set including all occurrences of taxa considered as putative NIS at at least one of our study locations. This allowed us to obtain information on the pan-European distribution of these taxa across their range of introduction as well as across their native range (for cases where the latter included one or more of our surveyed locations) (see Supplementary File 6). Given varying taxonomic rank specifications in reference databases, we manually corrected the taxonomy of putative NIS mOTUs to obtain correct phylum-level assignments. We then assessed phylum-wise relative abundances of the NIS in our data set, as well as the number of observatories each NIS occurred at. To quantify and visualise unique and shared occurrences of NIS across

observatories, an UpSet plot was generated. For this, we considered only the occurrences representing non-indigenous status at a given location (i.e., removing occurrences of a putative NIS at locations where it was considered native). Exploratory analyses were performed using *R* v4.3.1 and the packages *dplyr* v1.1.3 (Wickham, François, et al., 2023), *ggplot2* v3.5.1 (Wickham 2016), *scales* v1.3.0 (Wickham et al. 2023a, b, c), *forcats* v1.0.0 (Wickham 2023), *tidyr* v1.3.0 (Wickham et al. 2024), *RColorBrewer* v1.1.3 (Neuwirth 2014), *ggpubr* v0.6.0 (Kassambara 2023), *UpSetR* v1.4.0 (Conway et al. 2017), and *phyloseq* v1.44.0. All code utilised for analyses detailed below can be found at https://github.com/JustinePa/studyNIS_ARMS.

We then compared the occurrences of all detected NIS with their currently known range according to the information obtained during the manual curation, to identify potential range shifts (for details see Supplementary Text S2). For NIS that were found in previously unknown locations, we used the R package *rgbif* v3.8.0 (Chamberlain et al., 2024) to retrieve and plot their occurrences data from the Global Biodiversity Information Facility (GBIF) (GBIF 2024). We used GBIF occurrences as a proxy for the known range of each species (including both native and introduced ranges) and mapped occurrences deposited in GBIF and occurrences from our data set using the R package *rnaturalearth* v1.0.1 (Massicotte and South 2024) to visualise potential range shifts. See Supplementary File 7 for information on retrieved GBIF occurrences for each investigated NIS.

Standardisation of data sets and comparative analysis

We used the most unfiltered data set for a comprehensive assessment. However, to compare NIS prevalence between sites and monitoring areas, we ensured consistent sampling and sequencing efforts across data sets (see Supplementary Text S3).

We analysed data from both marker genes separately, removed certain replicate samples, and rarefied them to ensure even sequencing depth. We then again only considered the occurrences representing non-indigenous status at a given location (i.e., removing occurrences of a putative NIS at locations where it was considered native). Using the Kruskal–Wallis test in R, we assessed whether NIS prevalence was higher in areas impacted by maritime traffic (e.g., marinas, harbours, ports) compared to other areas. However, limited and unbalanced data across observatories prevented robust intra-observatory or intra-regional comparisons (Supplementary File 1). Future ARMS-MBON expansions may enable such analyses.

To standardise the data set, we accounted for variations in the number of sampled locations and deployed ARMS units across observatories by grouping ARMS units into "sites" based on spatial proximity. Using the *R* package *geosphere* v1.5.18 (Hijmans 2022), we calculated the distances between deployed ARMS units and grouped those with a maximum distance of 20 km from each other into a single "site." This grouping allowed us to standardise the terminology and contrast "sites" from "observatories." Since the number of ARMS units varied between sites, we randomly sub-sampled ARMS units within each site to equalise the sampling effort, ensuring comparability across sites. For comparability, we generated five "fraction" data sets for each marker gene based on the available size fractions after bioinformatic processing: (i) all three fractions, (ii) both motile fractions, (iii) sessile fraction, (iv) motile fraction 100 µm, and (v) motile fraction 500 µm. This allowed us to compare NIS counts across sites and fractions under equal sampling and sequencing conditions. The number of detected NIS per site and fraction was visualised using *QGIS 3.34* (Supplementary File 8).

Results

General description of the sequencing data set

After sequence filtering, denoising and clustering, the COI data set used for comprehensive NIS screening comprised 305 remaining samples with a total of 7,711,978 reads in 10,646 mOTUs composed of 24,317 ASVs. The 18S data set comprised 343 remaining samples containing a total of 14,722,467 reads in 13,771 mOTUs composed of 18,440 ASVs. Only around 27% of COI mOTUs could be classified at the phylum level, but these accounted for ~ 78% of the total read number. In the case of 18S, ~ 47% of mOTUs representing ~ 85% of all reads could be classified at the phylum level. For COI, 807 mOTUs (~ 8%) making up ~ 58% of

all reads could be assigned a species name. Classification to species level could be obtained for 1,076 18S mOTUs (~8%) which amounted to ~34% of the total read number.

Curation of putative NIS identification

We found 842 species registered in WRiMS for the larger marine regions covered in our study. Screening against this list, we obtained species matches for 78 COI mOTUs (equalling 78 species) containing 447 ASVs and for 35 mOTUs of 18S rRNA (equalling 35 species) comprising 154 ASVs (see Supplementary File 4).

We excluded 93 ASVs from the COI data set due to ambiguous classification or insufficient taxonomic resolution, and four ASVs that also matched native species. The exclusion of these 97 ASVs led to the removal of nine COI mOTUs. Additionally, two mOTUs matched with reference sequences for both *Watersipora subtorquata* and *W. subatra* and these were separately kept as *Watersipora* sp. as both species are considered NIS at their location of occurrence in our data set.

For 18S, 71 ASVs displayed ambiguous taxonomic assignments, and it could not be clearly determined if they represented a NIS, hence, these were discarded. The exclusion of these ASVs led to the elimination of ten mOTUs. Five mOTUs were manually assigned to a higher taxonomic rank as they were assigned to more than one taxon considered as NIS at the location of occurrence, and it could not be determined which species they belonged to.

We reviewed the origin and introduction status of each putative NIS based on available literature. Consequently, 16 mOTUs (43 ASVs) were excluded from the COI data set, and 5 mOTUs (52 ASVs) from the 18S data set, due to insufficient evidence of non-indigenous status at any location. Details of the manual curation process and consulted references are provided in Supplementary File 5.

We detected one more low-abundance NIS mOTU for the COI data set and two more low-abundance NIS mOTUs for the 18S data set when no threshold of 5 reads per ASV occurrence was applied. Only the low-abundant COI mOTU could be considered a potential NIS at its locations of occurrence (Radashevsky et al. 2020). See Supplementary Text S1 for details on these three taxa.

Overall description of the curated NIS data set

After validation of taxonomic classification and putative NIS status, 52 species (i.e., 52 mOTUs with 307 ASVs) and 20 species (i.e., 20 mOTUs with 31 ASVs) remained as putative NIS in the COI and 18S data sets, respectively (see Supplementary Fig. S1). These represented 63 unique taxa which were putatively non-indigenous at one or more locations of their occurrence. Of these, 43 NIS were detected with the COI marker only, eleven with the 18S marker only, and nine were detected by both genes.

The majority of detected NIS were arthropods ($n=16$), followed by molluscs ($n=9$), chordates ($n=8$), and bryozoans ($n=8$) (Fig. 2a). We analysed all occurrences of taxa classified as putative NIS at any study location to assess changes in their known distribution. The most widespread NIS, based on occurrences across observatories (not implying NIS status at all sites), included *Amphibalanus improvisus* (14 observatories), *Bonnemaisonia hamifera* (12), *Corella* sp. (8), *Bugula neritina* (7), *Balanus trigonus* (7), *Acartia (Acanthacartia) tonsa* (7), and *Hydroides elegans* (7) (Fig. 2b). Detailed pan-European occurrences of all 63 NIS are provided in Supplementary File 6.

Of the 12 observatories with multiple sampling events at the same location, 11 showed repeated detections of NIS (Supplementary File 6). Notably, the Koster observatory in Sweden consistently identified five NIS (*Acartia (Acanthacartia) tonsa*, *Bonnemaisonia hamifera*, *Corella* sp., *Dasysiphonia japonica*, and *Fenestrulina delicia*) from 2018 to 2020 across its three deployment sites. Similarly, Limfjord, Denmark, detected five species (*Caprella mutica*, *Dasysiphonia japonica*, *Fenestrulina delicia*, *Fibrocapsa japonica*, and *Polydora cornuta*) consistently between June 2019 and November 2021. These repeated detections indicate the continuous presence of these NIS in the monitored ecosystems.

Prevalence of NIS across observatories

Each of the 19 investigated ARMS-MBON observatories displayed occurrences of taxa considered non-indigenous at the respective location or at any of the other study locations (Supplementary Fig. S2) and all showed occurrences of at least one taxon that was

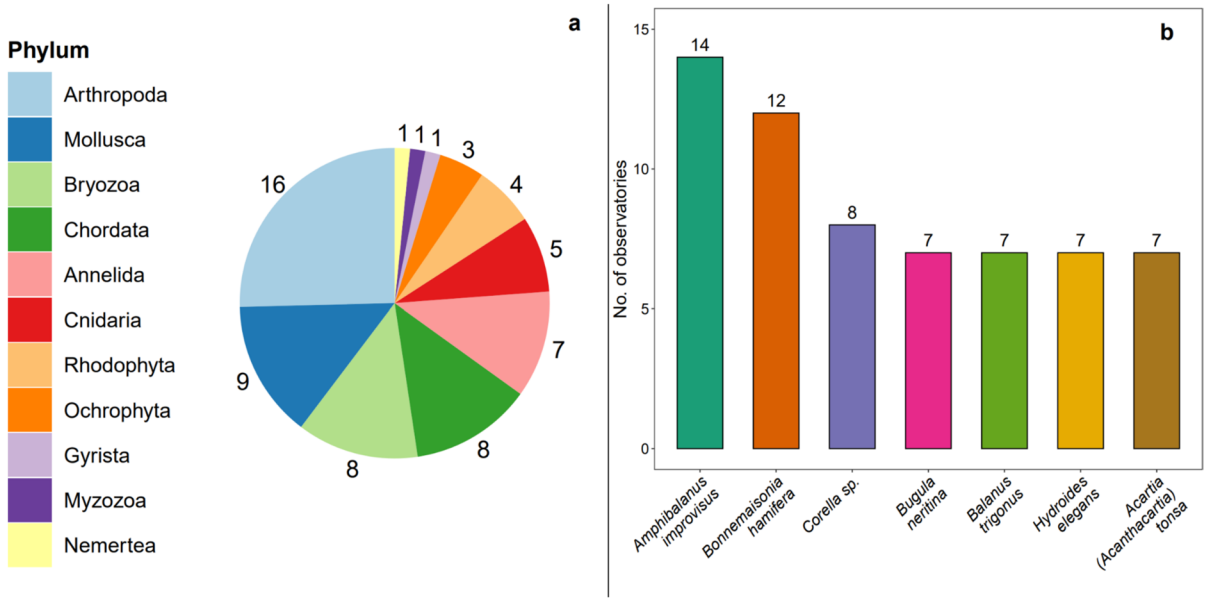

Fig. 2 **a** Taxonomic composition of the 63 identified NIS using the COI and 18S marker genes. Values represent the absolute number of NIS detected per phylum; **b** Bar chart depicting the most prevalent NIS based on the number of observatories they occurred at (number above each bar). Note that this does not imply putative NIS status at every observatory the respective taxon was present

considered as NIS in the respective area (Supplementary Fig. S3).

As shown in Fig. 3, 40 NIS were detected at more than one observatory while 23 NIS were detected at one observatory only. A similar UpSet plot displaying all occurrences of all NIS (i.e., occurrences across native ranges and locations of putative introduction) can be found in Supplementary Fig. S4.

The largest NIS sets were found in Plymouth, UK, and the Gulf of Piran, Slovenia, with 17 NIS at both locations, suggesting significant hubs for NIS. Four taxa were unique to Plymouth (*Boccardia proboscidea*, *Botrylloides violaceus*, *Schizoporella japonica* and *Watersipora* sp._3), while two were exclusive to Piran (*Asparagopsis taxiformis* and *Polycerella emertoni*), both known Mediterranean NIS.

RavennaM, Italy, showed a high number of detected NIS ($n=14$), five of which were unique to this observatory: *Cutleria multifida*, *Eucheilota menoni*, *Paracerceis sculpta*, *Haloa japonica*, and *Paranthura japonica.* All of them were previously known in the area, except *E. menoni*. In contrast, nearby RavennaH detected only 8 NIS, with *Mawia benovici* and *Xenostrobus securis* being unique to this site.

The Baltic Sea observatories (Gdynia, Poland, and TZS, Finland) detected fewer NIS ($n=6$ and $n=8$, respectively) including a number of taxa exclusively detected there. TZS had two unique taxa: *Dreissenidae* sp. (likely *Dreissena polymorpha* or *Mytilopsis leucophaeata*) and *Marenzellaria* sp. (reclassified as *Marenzellaria neglecta*). Both are known Baltic NIS. Gdynia uniquely detected *Pleurosira leavis*.

Detection and mapping of potential new introductions and range shifts

We detected several potential new introductions or range shifts, i.e. species in our curated NIS list that had occurrences in places where they had not been recorded previously according to GBIF and our literature search. For all these cases, occurrences were represented by a relatively low number of sequences reads.

Eucheilota menoni was detected on the north coast of the Adriatic Sea, at the RavennaM observatory in Italy, with 54 reads in the COI data set. At the observatory located in the Red Sea (Eilat, Israel), *Bugula neritina* was detected with 15 reads. *Fenestrulina delicia*, known on the Atlantic coasts of Europe, showed 8 reads in Crete, Greece, in the COI data set.

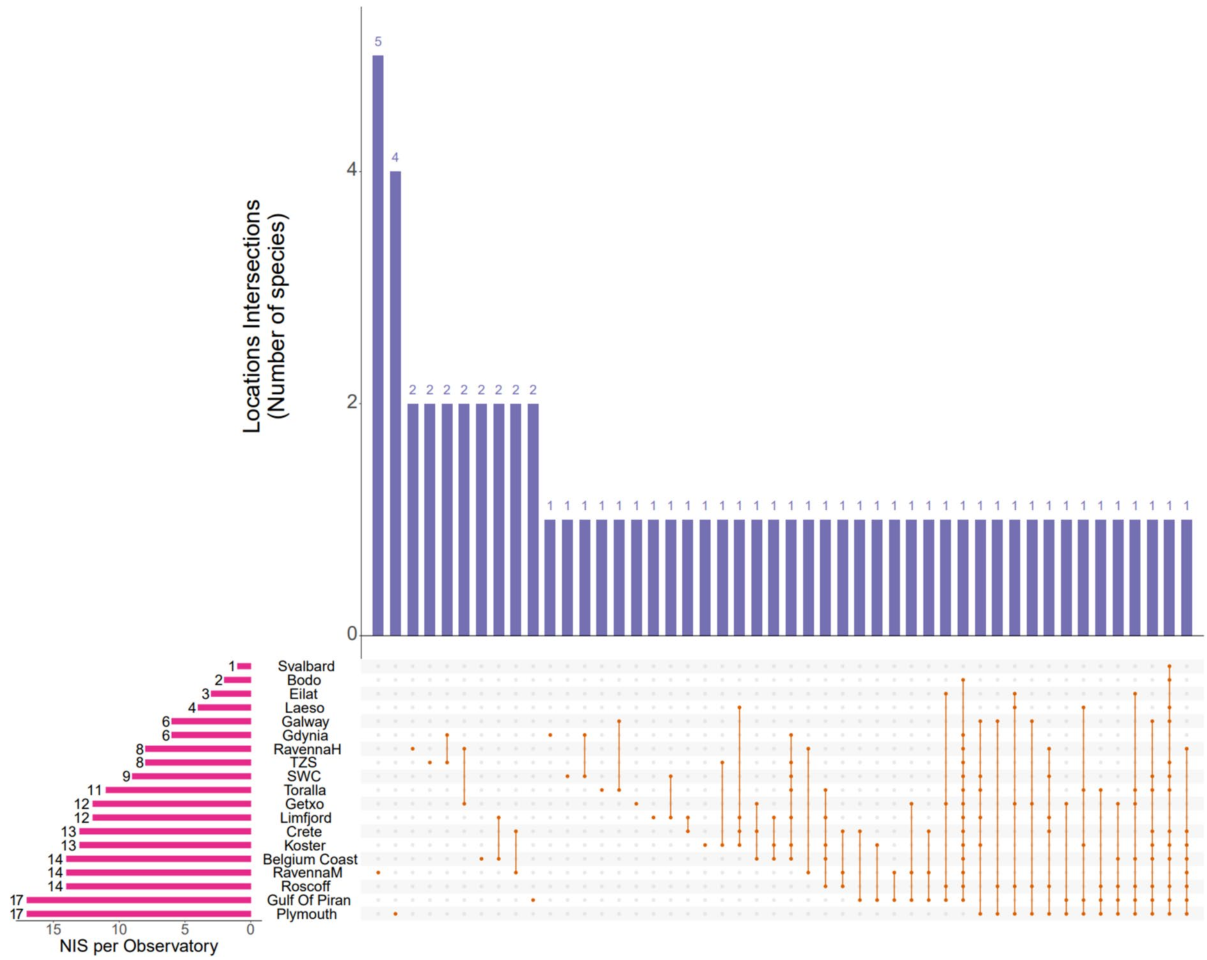

Fig. 3 UpSet plot showing the number of occurrences and co-occurrence of NIS across observatories. The upper vertical bar plot shows the number of NIS detected in any given combination of observatories; e.g., a specific set of five co-occurring NIS was only found at RavennaM and a set of four co-occurring NIS only at Plymouth, while there were eight cases of two NIS co-occurring at one or two observatories, etc. This only displays the occurrences representing non-indigenous status at a given location. The matrix below shows which sets of observatories are represented by each vertical bar. Importantly, each dot in the matrix below represents an occurrence at the respective observatory, but a single dot may correspond to multiple species. This explains why, for example, Plymouth shows fewer dots (in the matrix) than the total number of NIS, as some dots represent multiple NIS that co-occur across observatories. The horizontal bar plot on the left shows the number of taxa considered non-indigenous at each observatory

Additionally, a few species originating from regions with warmer water temperatures occurred with a low number of reads in observatories at higher latitudes. In Koster, Sweden, *Ostreopsis ovata* was found with 8 reads in the 18S data set. In the COI data set, *Apionsoma (Apionsoma) misakianum*, originating in the Indo-Pacific, was detected with 273 reads in Eilat, Israel, 10 reads in Getxo, Spain, 7 reads in Piran, Slovenia, and 7 reads in Roscoff, France. *Herdmania momus*, a known invader in the Mediterranean Sea from the Indo-Pacific Ocean, occurred with 6 reads in Limfjord, Denmark.

Mapping these novel detections against known occurrences from GBIF, we observed that they fell outside of the respective species' current known range (see examples in Fig. 4 and Supplementary Figs. S5–S9). Some were located further north compared to the current known distribution range (*e.g.*, *Ostreopsis ovata* or *Herdmania momus*), some were found in marine regions they had not been observed

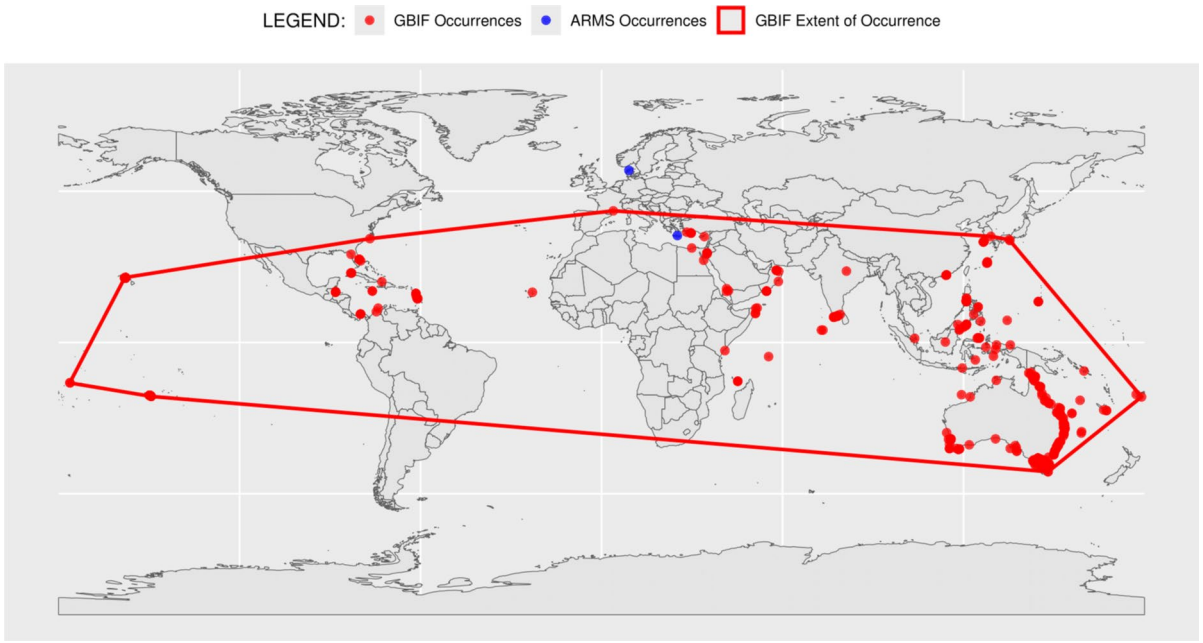

Fig. 4 Global extent of occurrences of *Herdmania momus* (in red) according to the GBIF database (red points), and the locations where this species has been detected in the present study (blue points). This species was detected in the COI data set in Crete, Greece, with 98 reads and in Limfjord, Denmark, with 6 reads

in before (e.g. *Fenestrulina delicia* found in the Mediterranean, but previously only known in the Atlantic).

Comparative analyses across observatories

Spatial distribution of non-indigenous species

Maps of NIS abundance (Fig. 5) revealed locations with high NIS abundance. The north of the Adriatic Sea appeared as the area with most NIS: Ravenna had 10.2 NIS per data set on average (Fig. 5a) and the Gulf of Piran had 2.8 NIS per data set on average (Fig. 5b). Figure 5a highlighted that Plymouth also hosted many NIS with an average of 5.4 per data set.

Detection of non-indigenous species in marinas and harbours

We tested for statistically significant differences in the number of taxa that could be considered NIS at their location of occurrence between types of deployment sites. Here, we compared the number of NIS (i.e., the number of taxa considered NIS at a specific location) detected in individual samples from locations associated with maritime traffic to the number of NIS in samples from all other locations. For the COI data set, after manual curation, on average 2.50 ± 2.08 NIS were detected in samples from ARMS deployed in marinas, harbours, and ports. Samples from all other locations displayed a mean number of 2.00 ± 1.51 detected NIS. Given that a lower number of putative NIS was recovered using 18S compared to COI data, the mean number of detected NIS per 18S sample was substantially lower. On average, 0.55 ± 0.83 NIS were detected in samples from ARMS deployed in marinas, harbours, and ports, while samples from all other locations recovered 0.50 ± 0.79 NIS. For both genes, differences between the two sample types were not statistically significant (COI: Kruskal–Wallis $\chi^2 = 1.273$, $p = 0.259$; 18S: Kruskal–Wallis $\chi^2 = 0.144$, $p = 0.705$). See Supplementary Table S2 for details on test results.

Discussion

Broad monitoring of NIS across regions, habitats, and taxa

In recent years, genetic methods have gained attention in the field of marine conservation and

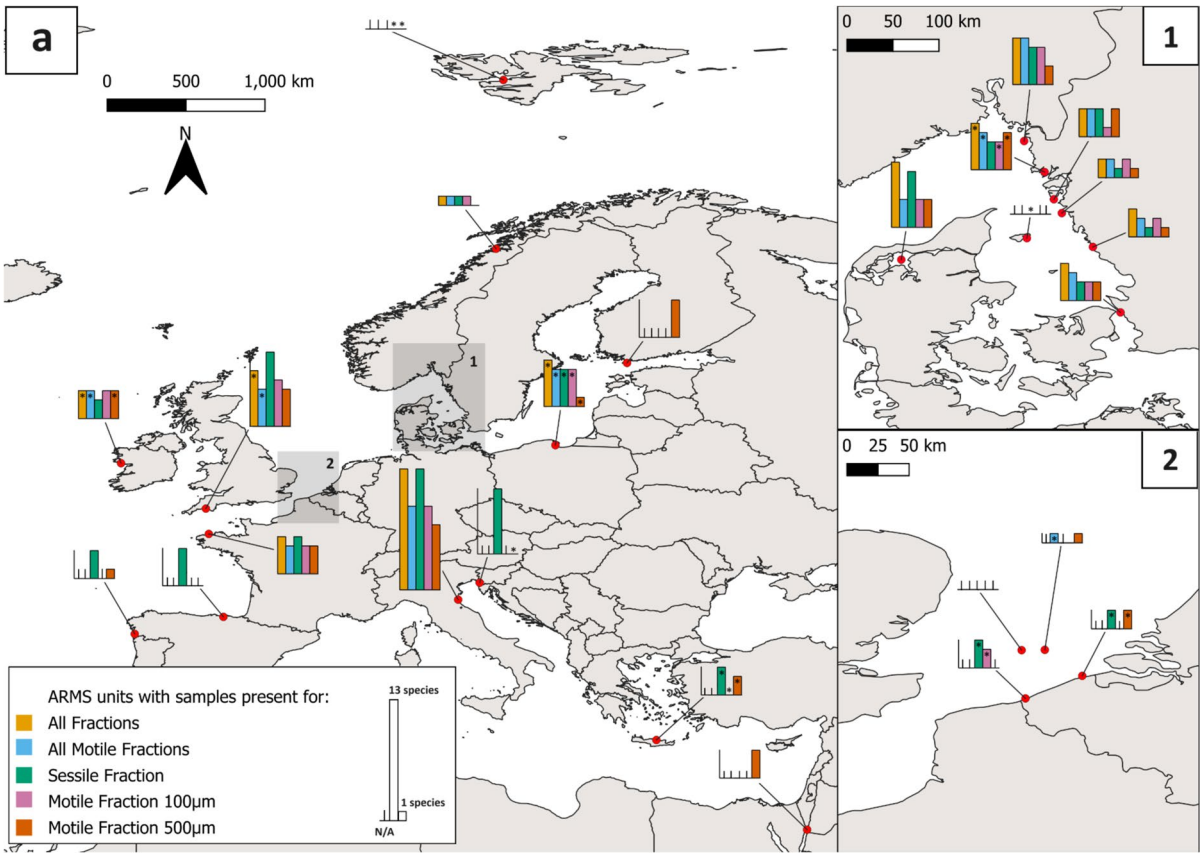

Fig. 5 Comparative distribution of NIS detected using COI (**a**) and 18S (**b**) molecular markers. Each bar graph shows the number of NIS per site across five data sets represented by distinct colours: ARMS with all fractions present (orange), ARMS with all motile fractions present (blue), ARMS with only the sessile fraction present (green), ARMS with only the motile 100 µm fractions present (pink), and ARMS with only the motile 500 µm fraction present (red), ensuring comparability across locations. Bars marked with an asterisk indicate data derived from a single ARMS replicate unit instead of two replicate units used for all other sites. Where data for a particular data set is absent (N/A), a tick mark is present instead of a bar to indicate no data was available for that category

management (Darling and Blum 2007; Westfall et al. 2020) and DNA metabarcoding has been promoted as a tool to monitor biological invasions (Comtet et al. 2015). While earlier surveillance efforts for aquatic NIS have primarily monitored one or a few species and have been limited geographically (Ficetola et al. 2008; Mahon et al. 2013; Tréguier et al. 2014), the methods used in this study allow for broad and continuous screening of NIS across the European regional seas, with sample locations in various habitats, and detecting a wide range of eukaryotic taxa.

During this study, 63 non-indigenous taxa were detected from 1,828 identified species across 112 ARMS-MBON sampling events in 19 observatories in 14 countries. Most of these NIS were previously known in the areas where they have been detected.

The most widespread NIS throughout the data set were *Amphibalanus improvisus*, *Bonnemaisonia hamifera, Bugula neritina, Balanus trigonus, Acartia tonsa* and *Hydroides elegans,* all well-documented marine invaders in European waters (details in Supplementary Text S4). Their widespread detection confirms that our approach reliably identifies established NIS and can monitor their future range expansion. This study highlights that a genetic monitoring network like ARMS-MBON, through continuous long-term deployments, is effective in detecting established NIS and offers valuable insights for ongoing marine management efforts.

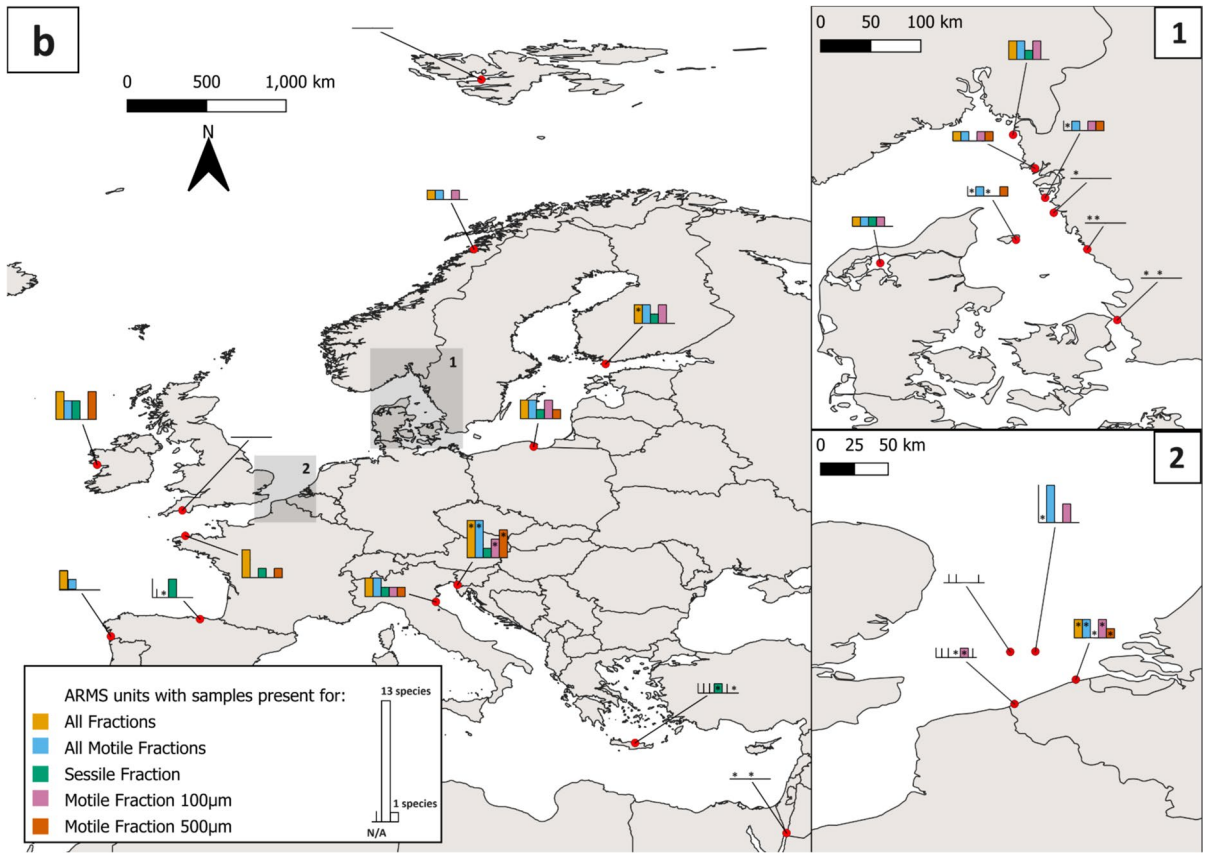

Fig. 5 (continued)

Early warning potential

Our study relied on WRiMS to screen for NIS, aiming to minimise false positives and negatives. While this approach enhanced result reliability, it limited the detection of newly introduced species to the 842 registered in WRiMS for our study regions. Despite these constraints, our large geographic scope enabled the identification of several potential new introductions, classified as such due to the absence of prior detections in the literature or other NIS databases (Supplementary Text S5).

Eucheilota menoni was detected on the north coast of the Adriatic Sea, in Ravenna, with 54 reads (Supplementary Fig. S9). We only detected it in our study as it is registered in WRiMS for both the Bay of Biscay and the Belgian coast where some of our observatories were located. However, we found no occurrences of this species on ARMS from observatories in these two regions.

Potential new invaders were also detected in the Red Sea. *Bugula neritina*, a colonial bryozoan of unknown origin, potentially from the Mediterranean Sea or the US, is widespread in Europe (Ryland et al. 2011) and serves as one example. No previous records of this species in the Red Sea were found in GBIF (Supplementary Fig. S6) or in OBIS, nor in WoRMS or in other scientific literature. Only one study from 1980 found it in the Suez Canal (Ghobashy and El Komy 1980). However, it appeared in our data set because *Bugula neritina* is recorded in WRiMS for four of our selected areas. It was detected with 15 reads during one sampling event in the Gulf of Aqaba.

Similarly, *Bougainvillia muscus,* which we detected manually and not based on WRiMS records, is also a potential new invader in the Red Sea. This species is a hydrozoan native and commonly reported in the Mediterranean Sea, the North Sea and English Channel (Lee II and Reusser,

2012), which was detected with a total of 25 reads in Eilat, Israel, in the COI data set (Supplementary Fig. S10).

Apionsoma (Apionsoma) misakianum was found beyond its native Indo-Pacific range, detected in Eilat, Slovenia, and the Atlantic coasts of Spain and France. It had previously been detected and considered non-indigenous in the Aegean Sea in 2007 (Açik 2008) and in other Mediterranean areas according to SeaLifeBase (Palomares and Pauly 2024), but in GBIF no record of this species can be found further north than the Red Sea (Supplementary Fig. S5). Our results suggest a potential range extension of this species outside of the Mediterranean Sea.

Fenestrulina delicia, a fouling bryozoan, showed 8 reads in Crete, Greece, in the COI data set. This species is known on the European Atlantic coasts from France to Norway (Supplementary Fig. S7) and could have been recently introduced in the Mediterranean (De Blauwe et al. 2014).

Ostreopsis ovata, a toxic dinoflagellate, coming from tropical regions was found with 8 reads in the 18S data set in the observatory located in Sweden (Supplementary Fig. S8). Granéli et al. 2011 concluded that current global warming can influence the geographical expansion to higher latitudes and biomass accumulation by blooms of *O. ovata*. It was previously known in the Mediterranean, on the Atlantic coasts of France and the Iberian Peninsula (Accoroni et al. 2024; Chomérat et al. 2022; David et al. 2013; Fabri-Ruiz et al. 2024).

Herdmania momus, a Lessepsian migrant from the Indo-Pacific (Evans et al. 2013; Gewing et al. 2014; Shenkar and Loya 2008), occurred with 6 reads in Limfjord, Denmark (Fig. 4). This species appears to be adapted to tropical climates; however, according to Gewing et al., (2019), the invasive population in the Mediterranean demonstrated significantly greater survivability under cold conditions than the native population from the Red Sea.

These results should be interpreted with caution, not only due to the relatively low read numbers with which these species were detected in the respective areas, but also because of the inherent limitations in taxonomic assignment. These detections, often based on low read numbers, could signify early introductions or sequencing artefacts. The reliability of species-level identification depends on genetic markers and reference databases. Further bioinformatic refinements and ARMS-MBON data will enhance confidence in these observations.

Some detections may reflect gaps in historical records rather than actual new introductions, particularly in under-sampled regions like the Red Sea. While our study focused on a three-year timeframe, it has already detected potential new introductions, demonstrating the early warning capabilities of DNA metabarcoding. However, given that many species lists, such as WRiMS, span longer timeframes and geographical regions, we recognize that certain species known to be present in Europe may not have been detected during this short-term survey. This does not undermine the effectiveness of metabarcoding as a surveillance tool, but rather emphasises the importance of long-term, continuous monitoring to align short-term detections with long-term data sets.

Mapping potential range shifts

By comparing our data set with GBIF records, which include many species already detected outside their native range, we identified discrepancies suggesting potential range expansions and recent introductions. The proximity of new detections to known occurrences provides insight into possible causes of these shifts. For instance, species detected slightly further north may reflect climate-driven expansion, while distant occurrences, such as European species in the Red Sea, could indicate shipping as a vector.

Scaling up the integration of recent sightings with established extent of occurrence data from databases like GBIF can help track range shifts, introduction pathways, and habitat suitability for NIS. This can also enable predictions of species' responses to various environmental conditions. As ARMS-MBON and global ecological data sets grow and adopt FAIR principles (Tanhua et al. 2019; Tenopir et al. 2020), this approach offers a robust framework for studying species distribution shifts. Combining these data with species distribution models and machine-learning algorithms can efficiently detect new NIS range shifts and predict their responses to changing environments.

Identification of potential NIS hotspots

We argue that large-scale monitoring networks like ARMS-MBON hold strong potential for studying NIS hotspots in the future. Here we tested here

the hypothesis that marinas and harbours tend to have higher NIS richness due to increased anthropogenic influence such as ballast water discharge (Bailey 2015) and the transport of fouling organisms (Ashton et al. 2022; Bailey et al. 2020; Chan et al. 2022; Georgiades et al. 2021). Although the three observatories with the highest NIS diversity (Ravenna, Gulf of Piran, Plymouth) were located near marinas or ports, we did not find a significantly higher number of NIS in samples from marinas and harbours compared to other locations in our data set. This suggests that additional environmental factors may also play a role in shaping NIS distribution.

Climate-driven factors like water temperature, salinity changes (Floerl et al. 2013; Occhipinti-Ambrogi 2021), and ocean currents, which facilitate the dispersal of species with planktonic larvae or rafting capabilities (e.g., *Hydroides elegans*, *Amphibalanus modestus*, *Bugula neritina*, *Watersipora* sp.), may contribute significantly to NIS spread. Thus, while marinas and harbours are critical entry points, they may not always coincide with the highest NIS prevalence. This highlights the complex interplay between anthropogenic and environmental drivers in shaping NIS distribution. Large-scale observatory networks such as ARMS MBON, when combined with environmental data, can help model these dynamics and improve predictions of NIS hotspots over time.

Methodological challenges

This study demonstrates the potential of our chosen methods for monitoring NIS across extensive geographical areas like the European coastline. We advocate for the increased use of DNA metabarcoding to detect and track NIS, but also for the application of this method at a large geographic scale, to have a better understanding of species migrations beyond the boundaries of countries or single marine regions. However, we also acknowledge various technical and practical challenges associated with this approach and emphasise the need for further methodological development. Each of the following challenges is further developed in Supplementary Text S6.

Challenges with large-scale marine biological monitoring and standardization

Recently, the concept of Essential Biodiversity Variables (EBVs) was developed to structure biodiversity monitoring at larger geographic scales and thereby enable continuous assessments of status and change of biodiversity as well as the driving forces behind these changes (Kissling et al 2018). EBVs critically depend on the implementation of consistent and comparable data collection campaigns across diverse ecosystems and regions (Canonico et al. 2019; Patrício et al. 2016). Here we showed that the standardised sampling protocols can be used to collect biodiversity variables for invasive species monitoring and management including tracking of non-indigenous species distributions and early detection of invasive species across borders (Lehtiniemi et al. 2015). Nevertheless, genetic monitoring demands transparent management of large and complex data according to FAIR data principles (Tanhua et al. 2019) as well as access to taxonomic expertise for manual curation steps. These resources need to be developed alongside the monitoring network and its sampling protocols. In addition, sampling needs to be intensified to cover more habitats and regions to deliver EBVs for non-indigenous species monitoring in the future.

Challenges in Taxonomic Identifications through DNA Metabarcoding

DNA metabarcoding is an effective tool for biodiversity monitoring but faces several limitations that impact its utility, particularly for the detection of NIS (Ruppert et al. 2019). One of the primary challenges is the incompleteness of reference databases and checklists, which often lack critical entries for taxa commonly identified as potential invaders (Hestetun et al. 2020; Chorlton 2024). In this study, many molecular operational taxonomic units (mOTUs) were only classifiable at higher taxonomic levels, with only 8% reaching species level, consistent with other studies (Pearman et al. 2021). This limits accurate NIS identification and contributes to knowledge gaps that affect management decisions and predictive models (Golo et al. 2023). Global initiatives like BOLD and GenBank are essential for this purpose, but are currently insufficient (Holman et al. 2019). Fontes et al. (2021) found in their study that only 42%

of NIS had records in BOLD, with significant discrepancies in over half of these records.

A related challenge is primer specificity. Universal primers, such as those targeting COI and 18S, are not truly universal and may introduce biases by amplifying certain taxa while excluding others (Jeunen et al. 2019). Multi-marker strategies have been recommended to improve taxonomic coverage (Portas et al. 2022) and provide more accurate community composition estimates (Günther et al. 2018; Stefanni et al. 2018). However, this approach requires careful optimization to avoid underrepresentation of key taxa. Additionally, suboptimal protocols and the transient nature of environmental DNA further complicate detection, potentially leading to false negatives (Collins et al. 2018).

Incomplete sampling across habitats, space, and time also presents significant challenges. ARMS, for instance, target specific habitats and may miss species that do not settle on artificial structures (Zhan et al. 2014). Temporal limitations can further skew results; since our data set spans only three years, natural fluctuations or delayed introductions might explain the absence of certain species.

Bioinformatic errors in taxonomic assignment introduce another layer of complexity. Clustering methods can misidentify closely related species or split a single species into multiple mOTUs (Brown et al. 2016) and manual curation of each NIS mOTU is needed to minimise such errors (Antich et al. 2021). Continued refinement of bioinformatics pipelines, coupled with manual validation, is crucial for accurate species identification.

Contamination remains an inherent risk, particularly in low-biomass samples. Laboratory contamination or cross-sample contamination can introduce false positives, hence the need for a curation. Rigorous protocols, including the use of negative controls and decontamination procedures, are essential to mitigate this risk.

Despite these limitations, metabarcoding offers many advantages over conventional NIS surveys, which are also prone to errors such as observer bias and under-detection of cryptic species. Metabarcoding provides broader taxonomic coverage and can reveal hidden biodiversity, but it is inherently semi-quantitative. The number of reads does not directly correlate with species abundance, emphasising the need for presence-absence analysis. Complementing metabarcoding with traditional morphological assessments and visual surveys can enhance monitoring accuracy (Cowart et al. 2015; Thomas et al. 2016; Yu et al. 2012). Programs like EMO BON (Santi et al. 2023), which integrate metabarcoding of soft-bottom habitats and water columns, demonstrate the value of a multi-faceted approach.

Challenges with NIS databases and the importance of manual curation

Multiple databases record the introduction status of marine species, and for this study, we relied on WRiMS. However, WRiMS, like all current databases, is neither fully curated nor comprehensive. Our comparison of WRiMS with GRIIS, marInvaders, and AquaNIS (Supplementary Figs. S11 and S12) highlights that combining databases provides the most comprehensive coverage.

Manual curation is essential in any metabarcoding study (Deagle et al. 2018; Westfall et al. 2020). In our study, manual curation filtered out around 40% of potential NIS misclassifications. For instance, *Carcinus maenas*, native to European waters, was flagged as an NIS due to its potential introduction to the Red Sea in WRiMS, though it was not detected in our Red Sea observatory and was thus excluded (Roman et al., 2004). Conversely, *Juxtacribrilina mutabilis*, an expanding bryozoan, was detected in several observatories but is not listed in WRiMS for Europe, revealing a database gap (Ito et al. 2015; Martaeng et al. 2023; Dick et al. 2020). Similarly, *Bougainvillia muscus* was detected in the Red Sea but was absent from WRiMS for this region. Manual literature searches confirmed its introduction status, demonstrating the limitations of relying solely on databases.

Despite the limitations of current NIS databases, including WRiMS, we acknowledge these challenges and caution future users to approach them critically. However, by employing a rigorous manual curation process and cross-referencing multiple databases, accuracy and reliability can be ensured.

Conclusion

This study investigated the potential of a genetic observatory network to detect and monitor marine non-indigenous species and advance our

understanding of such taxa, which is crucial to preserve biodiversity and maintain healthy ecosystems. By identifying 63 non-indigenous species, we showed that genetic observatory networks can monitor distant phylogenetic groups across a vast geographical region and detect potential new introductions and range shifts at the earliest possible stage. We also showed the need for a manual curation to obtain reliable results. We highlight the potential of these networks to identify areas with high NIS abundance and habitats vulnerable to NIS establishment.

Conflict of interest

All authors certify that they have no affiliations with or involvement in any organisation or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Acknowledgements We thank the numerous researchers, field workers and technical staff across the ARMS MBON network for their dedicated efforts in data collection and logistical support. We appreciate the reviewers and Editor-in-Chief for their contributions to the peer review process. We thank Katrina Exter and Christina Pavloudi for their support with managing and accessing the data. Thanks to all the staff of the Laboratory of Ecological and Environmental Genetics (GEA), Department of Biological, Geological and Environmental Science, University of Bologna for the installation and construction of the ARMS in Ravenna. This work used resources provided by the European Marine Omics Biodiversity Observation Network (EMO BON) project, coordinated by the European Marine Biological Resource Centre (EMBRC). We acknowledge the support of various funding organisations and projects: ASSEMBLE Plus (grant no. 730984), DTO-bioflow (grant no. 101112823) and MARCO BOLO (grant no. 101082021), the INTERREG GEANS project, the Swedish Agency for Marine and Water Management (grant no. 3181-2019), the Flanders LifeWatch contribution (Research Foundation Flanders grant I000819N), and the Aquanis 2.0 project (FONDATION Total). We also thank the Swedish Biodiversity Data Infrastructure (grant no. 2019-00242), the IMBBC of the HCMR, MARBIGEN, LifeWatchGreece RI, and the CMBR RI for data management and computational resources. The computations and data processing were also enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS), partially funded by the Swedish Research Council through grant agreement no. 2022-06725

Author contributions All authors contributed to the study conception and design. Material preparation and data collection were executed by JP and ND, with ND handling the processing of sequencing data and JP performing the manual curation of the initial NIS data set. Data analysis was performed jointly by JP and ND. The first draft of the manuscript was written by JP, with substantial input and feedback from ND. MO provided critical revisions and feedback on all versions of the manuscript. MO initiated and is leading ARMS-MBON, from which the data was obtained, and provided resources and supervision. All authors read and approved the final manuscript.

Funding Open access funding provided by University of Gothenburg. The ARMS-MBON network was established under the infrastructure program ASSEMBLE Plus (grant no. 730984) and is currently financed by the European Marine Biological Resource Centre (EMBRC) under the program European Marine Omics Biodiversity Observation Network (EMO BON). Data publication and analysis is funded by the projects DTO-bioflow (grant agreement no. 101112823) and MARCO BOLO (grant agreement no. 101082021). Funding for individual ARMS observatories was provided by the INTERREG project GEANS (North Sea Program of the European Regional Development Fund of the European Union), the Swedish Agency for Marine and Water Management (grant no. 3181-2019), the Flanders LifeWatch contribution (Research Foundation Flanders grant I000819N), and the Aquanis 2.0 project (FONDATION Total). Data management and computational resources were provided by the Swedish Biodiversity Data Infrastructure (grant no. 2019-00242) and by IMBBC (Institute of Marine Biology, Biotechnology and Aquaculture) of the HCMR (Hellenic Centre of Marine Research). Funding for establishing the IMBBC HPC has been received by the MARBIGEN (EU Regpot) project, LifeWatchGreece RI and the CMBR (Center for the study and sustainable exploitation of Marine Biological Resources) RI.

Data availability All raw ARMS MBON data and metadata supporting this work are accessible through the ARMS MBON GitHub repository at https://github.com/arms-mbon. Standard operating procedures and protocols are available on the dedicated ARMS-MBON GitHub repository (https://github.com/armsmbon/documentation). All metadata and access to all image data generated by ARMS-MBON to date can be found on GitHub (https://github.com/arms-mbon/data_workspace). All genetic raw data generated by ARMS-MBON to date can be accessed on the European Nucleotide Archive (ENA) through the accession numbers provided via the GitHub repository (https://github.com/arms-mbon/data_workspace) and under the umbrella study PRJEB72316 (https://www.ebi.ac.uk/ena/browser/view/PRJEB72316). The code for the bioinformatic pipeline and the initial NIS scan is published in Daraghmeh (2024). The code used for data exploration and visualisation can be found at https://github.com/JustinePa/studyNIS_ARMS.

References

Accoroni S, Neri F, Ubaldi M, Romagnoli T, Totti C (2024) *Ostreopsis* cf. ovata abundances on different benthic substrata: how to compare them? Marine Biol. https://doi.org/10.1007/s00227-023-04328-8

Açik S (2008) Occurrence of the alien species *Aspidosiphon* (*Aspidosiphon*) *elegans* (*Sipuncula*) on the Levantine and Aegean coasts of Turkey. Turkish J Zool 32(4):443–448

Ahyong S, Boyko CB, Bailly N, Bernot J, Bieler R, Brandão SN, Daly M, De Grave S, Gofas S, Hernandez F, Hughes L, Neubauer TA, Paulay G, Boydens B, Decock W, Dekeyzer S, Goharimanesh M, Vandepitte L, Vanhoorne B, Zullini A (2024). World Register of Marine Species (WoRMS). WoRMS Editorial Board

Antich A, Palacin C, Wangensteen OS, Turon X (2021) To denoise or to cluster, that is not the question: optimizing pipelines for COI metabarcoding and metaphylogeography. BMC Bioinf 22(1):1–24. https://doi.org/10.1186/S12859-021-04115-6

AquaNIS. Editorial Board, (2015) Information system on Aquatic Non-Indigenous and Cryptogenic Species. World Wide Web electronic publication. www.corpi.ku.lt/databases/aquanis. Version 2.36+. Accessed 2024–10–21.

Ashton GV, Zabin CJ, Davidson IC, Ruiz GM (2022) Recreational boats routinely transfer organisms and promote marine bioinvasions. Biol Invasions 24(4):1083–1096. https://doi.org/10.1007/s10530-021-02699-x

Atienza S, Guardiola M, Præbel K, Antich A, Turon X, Wangensteen OS (2020) DNA metabarcoding of Deep-Sea sediment communities using COI: community assessment, spatio-temporal patterns and comparison with 18S rDNA. Diversit 12(4):123. https://doi.org/10.3390/D12040123

Bailey SA (2015) An overview of thirty years of research on ballast water as a vector for aquatic invasive species to freshwater and marine environments. Aquat Ecosyst Health Manage 18(3):261–268. https://doi.org/10.1080/14634988.2015.1027129

Bailey SA, Brown L, Campbell ML, Canning-Clode J, Carlton JT, Castro N, Chainho P, Chan FT, Creed JC, Curd A, Darling J, Fofonoff P, Galil BS, Hewitt CL, Inglis GJ, Keith I, Mandrak NE, Marchini A, McKenzie CH, Zhan A (2020) Trends in the detection of aquatic non-indigenous species across global marine, estuarine and freshwater ecosystems: a 50-year perspective. Divers Distribut 26(12):1780–1797

Bakker J, Wangensteen OS, Baillie C, Buddo D, Chapman DD, Gallagher AJ, Guttridge TL, Hertler H, Mariani S (2019) Biodiversity assessment of tropical shelf eukaryotic communities via pelagic eDNA metabarcoding. Ecol Evol 9(24):14341–14355. https://doi.org/10.1002/ECE3.5871

Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2012) GenBank. Nucleic Acids Res 41(D1):D36–D42. https://doi.org/10.1093/nar/gks1195

Bork P, Bowler C, De Vargas C, Gorsky G, Audic S, Henry N, Decelle J, Mahé F, Logares R, Lara E, Berney Ć, Le Bescot N, Probert I, Carmichael M, Poulain J, Romac S, Colin S, Aury JM, Bittner L, Velayoudon D (2015) Eukaryotic plankton diversity in the sunlit ocean. Science. https://doi.org/10.1126/science.1261605

Borrell YJ, Miralles L, Do Huu H, Mohammed-Geba K, Garcia-Vazquez E (2017) DNA in a bottle—rapid metabarcoding survey for early alerts of invasive species in ports. PLoS ONE. https://doi.org/10.1371/journal.pone.0183347

Brown EA, Chain FJJ, Zhan A, MacIsaac HJ, Cristescu ME (2016) Early detection of aquatic invaders using metabarcoding reveals a high number of non-indigenous species in Canadian ports. Divers Distrib 22(10):1045–1059. https://doi.org/10.1111/ddi.12465

Buchner D, Leese F (2020) BOLDigger—a Python package to identify and organise sequences with the barcode of life data systems. Metabarcod Metagenom 4:19–21. https://doi.org/10.3897/mbmg.4.53535

BJCallahan2018Silva taxonomic training data formatted for DADA2 (Silva version 132)Zenodo10.5281/zenodo.1172783Callahan BJ (2018) Silva taxonomic training data formatted for DADA2 (Silva version 132). Zenodo. https://doi.org/10.5281/zenodo.1172783

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: high-resolution sample inference from Illumina amplicon data. Nat Methods 13(7):581–583. https://doi.org/10.1038/nmeth.3869

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinf 10(1):1–9. https://doi.org/10.1186/1471-2105-10-421/FIGURES/4

Canonico G, Buttigieg PL, Montes E, Muller-Karger FE, Stepien C, Wright D, Benson A, Helmuth B, Costello M, Sousa-Pinto I, Saeedi H, Newton J, Appeltans W, Bednaršek N, Bodrossy L, Best BD, Brandt A, Goodwin KD, Iken K, Murton B (2019) Global observational needs and resources for marine biodiversity. Front Marine Sci. https://doi.org/10.3389/fmars.2019.00367

Catlett D, Son K, Liang C (2023) Ensembletax: ensemble taxonomic assignments of amplicon sequencing data

Cecchetto M, Dettai A, Gallut C, Obst M, Kuklinski P, Balazy P, Chelchowski M, Małachowicz M, Poćwierz-Kotus A, Zbawicka M, Reiss H, Eléaume MP, Ficetola GF, Pavloudi C, Exter K, Fontaneto D, Schiaparelli S (2024) Seasonality of primary production explains the richness of pioneering benthic communities. Nat Commun 15:8340. https://doi.org/10.1038/s41467-024-52673-z

Chan FT, Ogilvie D, Sylvester F, Bailey SA (2022) Ship biofouling as a vector for non-indigenous aquatic species to Canadian Arctic Coastal ecosystems: a survey and modeling-based assessment. Front Marine Sci. https://doi.org/10.3389/fmars.2022.808055

Charif D, Lobry JR (2007) Seqin{R} 1.0–2: a contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. In: Bastolla U, Porto M, Roman HE, Vendruscolo M (eds)

Structural approaches to sequence evolution: Molecules, networks, populations. Springer Verlag, pp 207-232.

Chomérat N, Antajan E, Auby I, Bilien G, Carpentier L, de Casamajor MN, Ganthy F, Hervé F, Labadie M, Méteigner C, Paradis C, Perrière-Rumèbe M, Sanchez F, Séchet V, Amzil Z (2022) First characterization of Ostreopsis cf. ovata (*Dinophyceae*) and detection of ovatoxins during a multispecific and toxic ostreopsis bloom on french atlantic coast. Mar Drugs 20(7):461. https://doi.org/10.3390/md20070461

Chorlton SD (2024) Ten common issues with reference sequence databases and how to mitigate them. Front Bioinf. https://doi.org/10.3389/fbinf.2024.1278228

Collins RA, Wangensteen OS, O'Gorman EJ, Mariani S, Sims DW, Genner MJ (2018) Persistence of environmental DNA in marine systems. Commun Biol. https://doi.org/10.1038/s42003-018-0192-6

Comtet T, Sandionigi A, Viard F, Casiraghi M (2015) DNA (meta) barcoding of biological invasions: a powerful tool to elucidate invasion processes and help managing aliens. Biol Invas 17(3):905–922. https://doi.org/10.1007/s10530-015-0854-y

Conway JR, Lex A, Gehlenborg N (2017) UpSetR: an R package for the visualization of intersecting sets and their properties. Bioinformatics 33(18):2938–2940. https://doi.org/10.1093/bioinformatics/btx364

Costello MJ, Dekeyzer S, Galil BS, Hutchings P, Katsanevakis S, Pagad S, Robinson TB, Turon X, Vandepitte L, Vanhoorne B, Verfaille K, Willan RC, Rius M (2021) Introducing the world register of introduced marine species (WRiMS). Manag Biol Invas 12(4):792–811. https://doi.org/10.3391/MBI.2021.12.4.02

Costello MJ, Ahyong S, Bieler R, Boudouresque C, Desiderato A, Downey R, Galil BS, Gollasch S, Hutchings P, Kamburska L, Katsanevakis S, Kupriyanova E, Lejeusne C, Ma KCK, Marchini A, Occhipinti A, Pagad S, Panov VE, Poore GCB, Zhan A (2024) World Register of Introduced Marine Species (WRiMS). WoRMS Editorial Board

Coster SS, Dillon MN, Moore W, Merovich GT (2021) The update and optimization of an eDNA assay to detect the invasive rusty crayfish (*Faxonius rusticus*). PLoS ONE. https://doi.org/10.1371/journal.pone.0259084

Couton M, Comtet T, Le Cam S, Corre E, Viard F (2019) Metabarcoding on planktonic larval stages: an efficient approach for detecting and investigating life cycle dynamics of benthic aliens. Manag Biol Invas 10(4):657–689. https://doi.org/10.3391/mbi.2019.10.4.06

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

Creedy TJ, Andújar C, Meramveliotakis E, Noguerales V, Overcast I, Papadopoulou A, Morlon H, Vogler AP, Emerson BC, Arribas P (2022) Coming of age for COI metabarcoding of whole organism community DNA: towards bioinformatic harmonisation. Mol Ecol Res. https://doi.org/10.1111/1755-0998.13502

Csárdi G, Hester J, Wickham H, Chang W, Morgan M, Tenenbaum D (2024) Remotes: R Package Installation from Remote Repositories, Including "GitHub"

Daraghmeh N, Exter K, Pagnier J, Balazy P, Cancio I, Chatzigeorgiou G, Chatzinikolaou E, Chelchowski M, Chrismas NAM, Comtet T, Dailianis T, Deneudt K, Diaz De Cerio O, Digenis M, Gerovasileiou V, Gonzalez J, Kauppi L, Kristoffersen JB, Kuklinski P, Obst M (2024) A long-term ecological research data set from the marine genetic monitoring programme ARMS-MBON 2018–2020. *bioRxiv,* https://doi.org/10.1101/2024.09.26.614897

Daraghmeh N (2024) ARMS-MBON 18S rRNA and COI gene metabarcoding: scanning for non-indigenous species. *protocols.io* https://doi.org/10.17504/protocols.io.n92ldmmmnl5b/v1

Darling JA, Blum MJ (2007) DNA-based methods for monitoring invasive species: a review and prospectus. Biol Invas. https://doi.org/10.1007/s10530-006-9079-4

Darling JA, Mahon AR (2011) From molecules to management: adopting DNA-based methods for monitoring biological invasions in aquatic environments. Environ Res 111(7):978–988. https://doi.org/10.1016/j.envres.2011.02.001

David H, Laza-Martínez A, Miguel I, Orive E (2013) *Ostreopsis* cf. siamensis and *Ostreopsis* cf. ovata from the Atlantic Iberian Peninsula: morphological and phylogenetic characterization. Harmful Algae 30:44–55. https://doi.org/10.1016/j.hal.2013.08.006

Davis TL (2023) Argparse: command line optional and positional argument parser. https://cran.r-project.org/package=argparse

De Blauwe H, Kind B, Kuhlenkamp R, Cuperus J, van der Weide B, Kerckhof F (2014) Recent observations of the introduced Fenestrulina delicia Winston, Hayward and Craig, 2000 (Bryozoa) in Western Europe. Studi Trent Sci Nat 94:45–51

Deagle BE, Clarke LJ, Kitchener JA, Polanowski AM, Davidson AT (2018) Genetic monitoring of open ocean biodiversity: an evaluation of DNA metabarcoding for processing continuous plankton recorder samples. Mol Ecol Resour 18(3):391–406

Dejean T, Valentini A, Miquel C, Taberlet P, Bellemain E, Miaud C (2012) Improved detection of an alien invasive species through environmental DNA barcoding: the example of the American bullfrog *Lithobates catesbeianus*. J Appl Ecol 49(4):953–959. https://doi.org/10.1111/j.1365-2664.2012.02171.x

Dick MH, Waeschenbach A, Trott TJ, Onishi T, Beveridge C, Bishop JD, Ito M, Ostrovsky AN (2020) Global distribution and variation of the invasive cheilostome bryozoan cribrilina mutabilis. Zoolog Sci 37(3):217. https://doi.org/10.2108/zs190142

Dowle M, Srinivasan A (2021) Data.table: extension of `data.frame`. https://cran.r-project.org/package=data.table

Dragulescu A, Arendt C (2020) Xlsx: read, write, format excel 2007 and Excel 97/2000/XP/2003 Files. https://cran.r-project.org/package=xlsx

Evans J, Borg JA, Schembri PJ (2013) First record of Herdmania momus (*Ascidiacea*: *Pyuridae*) from the central Mediterranean Sea. Marine Biodivers Rec 6:e134. https://doi.org/10.1017/S1755267213001127

Fabri-Ruiz S, Berdalet E, Ulses C, Somot S, Vila M, Lemée R, Irisson JO (2024) Harmful *Ostreopsis* cf. ovata blooms

could extend in time span with climate change in the Western Mediterranean Sea. Sci Total Environ. https://doi.org/10.1016/j.scitotenv.2024.174726

Ficetola GF, Miaud C, Pompanon F, Taberlet P (2008) Species detection using environmental DNA from water samples. Biol Let 4(4):423–425. https://doi.org/10.1098/rsbl.2008.0118

Floerl O, Rickard G, Inglis G, Roulston H (2013) Predicted effects of climate change on potential sources of non-indigenous marine species. Divers Distrib 19(3):257–267. https://doi.org/10.1111/ddi.12048

Fonseca VG, Davison PI, Creach V, Stone D, Bass D, Tidbury HJ (2023) The application of eDNA for monitoring aquatic non-indigenous species: practical and policy considerations. Diversity. https://doi.org/10.3390/d15050631

Fontes JT, Vieira PE, Ekrem T, Soares P, Costa FO (2021) BAGS: an automated barcode, audit and grade system for DNA barcode reference libraries. Mol Ecol Resour 21(2):573–583. https://doi.org/10.1111/1755-0998.13262

Fraija-Fernández N, Bouquieaux MC, 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. Ecol Evol 10(14):7560–7584. https://doi.org/10.1002/ece3.6482

Frøslev TG, Kjøller R, Bruun HH, Ejrnæs R, Brunbjerg AK, Pietroni C, Hansen AJ (2017) Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. Nat Commun 8(1):1–11. https://doi.org/10.1038/s41467-017-01312-x

Gallage KS, Nynatten AV, Lujan NK, Lovejoy NR, Mandrak NE (2023) Identifying early life stages of Great Lakes fishes using a metabarcoding approach. Can J Fish Aquat Sci 80(11):1813–1823. https://doi.org/10.1139/cjfas-2023-0061

GBIF: The Global Biodiversity Information Facility (2024) What is GBIF? Available from https://www.gbif.org/what-is-gbif [7 November 2024].

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. Mol Ecol Resour 13(5):851–861. https://doi.org/10.1111/1755-0998.12138

Georgiades E, Scianni C, Davidson I, Tamburri MN, First MR, Ruiz G, Ellard K, Deveney M, Kluza D (2021) The role of vessel biofouling in the translocation of marine pathogens: management considerations and challenges. Front Marine Sci. https://doi.org/10.3389/fmars.2021.660125

Gewing MT, Rothman SBS, Nagar LR, Shenkar N (2014) Early stages of establishment of the non-indigenous ascidian herdmania momus (Savigny, 1816) in shallow and deep water environments on natural substrates in the mediterranean sea. BioInvasions Rec 3(2):77–81. https://doi.org/10.3391/bir.2014.3.2.04

Gewing MT, Goldstein E, Buba Y, Shenkar N (2019) Temperature resilience facilitates invasion success of the solitary ascidian Herdmania momus. Biol Invasions 21(2):349–361. https://doi.org/10.1007/s10530-018-1827-8

Ghobashy AFA, El Komy MM (1980) Fouling in the southern region of the Suez Canal. Hydrobiol Bull 14(2):179–185. https://doi.org/10.1007/BF02260119

Goldberg CS, Sepulveda A, Ray A, Baumgardt J, Waits LP (2013) Environmental DNA as a new method for early detection of New Zealand mudsnails (*Potamopyrgus antipodarum*). Freshwater Science 32(3):792–800. https://doi.org/10.1899/13-046.1

Goldberg CS, Turner CR, Deiner K, Klymus KE, Thomsen PF, Murphy MA, Spear SF, McKee A, Oyler-McCance SJ, Cornman RS, Laramie MB, Mahon AR, Lance RF, Pilliod DS, Strickler KM, Waits LP, Fremier AK, Takahara T, Herder JE, Taberlet P (2016) Critical considerations for the application of environmental DNA methods to detect aquatic species. Methods Ecol Evol. https://doi.org/10.1111/2041-210X.12595

Golo R, Vergés A, Díaz-Tapia P, Cebrian E (2023) Implications of taxonomic misidentification for future invasion predictions: evidence from one of the most harmful invasive marine algae. Marine Pollut Bull. https://doi.org/10.1016/j.marpolbul.2023.114970

Granéli E, Vidyarathna NK, Funari E, Cumaranatunga PRT, Scenati R (2011) Can increases in temperature stimulate blooms of the toxic benthic dinoflagellate Ostreopsis ovata? Harmful Algae 10(2):165–172. https://doi.org/10.1016/j.hal.2010.09.002

Grey EK, Bernatchez L, Cassey P, Deiner K, Deveney M, Howland KL, Lacoursière-Roussel A, Leong SCY, Li Y, Olds B, Pfrender ME, Prowse TAA, Renshaw MA, Lodge DM (2018) Effects of sampling effort on biodiversity patterns estimated from environmental DNA metabarcoding surveys. Sci Rep. https://doi.org/10.1038/s41598-018-27048-2

Guillou L, Bachar D, Audic S, Bass D, Berney C, Bittner L, Boutte C, Burgaud G, De Vargas C, Decelle J, Del Campo J, Dolan JR, Dunthorn M, Edvardsen B, Holzmann M, Kooistra WHCF, Lara E, Le Bescot N, Logares R, Christen R (2013) The protist ribosomal reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. Nucl Acids Res 41(D1):D597–D604. https://doi.org/10.1093/NAR/GKS1160

Günther B, Knebelsberger T, Neumann H, Laakmann S, Martínez Arbizu P (2018) Metabarcoding of marine environmental DNA based on mitochondrial and nuclear genes. Sci Rep. https://doi.org/10.1038/s41598-018-32917-x

Hablützel PI, Martínez PA, Bilsen A, Christodoulou M, Delacauw S, Deneudt K, Khodami S, Lagaisse R, Heynderickx H, Obst M, Sapkota R, Sundberg P, Uhlir C, Staehr PAU, Winding A (2023) DNA-based monitoring of Non-Indigenous Species (NIS)

Hardy CM, Krull ES, Hartley DM, Oliver RL (2010) Carbon source accounting for fish using combined DNA and stable isotope analyses in a regulated lowland river weir pool. Mol Ecol 19(1):197–212. https://doi.org/10.1111/J.1365-294X.2009.04411.X

Hatzenbuhler C, Kelly JR, Martinson J, Okum S, Pilgrim E (2017) Sensitivity and accuracy of high-throughput metabarcoding methods for early detection of invasive fish species. Sci Rep. https://doi.org/10.1038/srep46393

Hestetun JT, Bye-Ingebrigtsen E, Nilsson RH, Glover AG, Johansen PO, Dahlgren TG (2020) Significant taxon sampling gaps in DNA databases limit the operational use of marine macrofauna metabarcoding. Marine Biodivers. https://doi.org/10.1007/s12526-020-01093-5

Hijmans RJ (2022) Geosphere: spherical trigonometry. https://cran.r-project.org/package=geosphere

Hoffman JC, Meredith C, Pilgrim E, Trebitz A, Hatzenbuhler C, Kelly JR, Peterson G, Lietz J, Okum S, Martinson J (2021) Comparison of larval fish detections using morphology-based taxonomy versus high-throughput sequencing for invasive species early detection. Can J Fish Aquat Sci 78(6):752–764. https://doi.org/10.1139/cjfas-2020-0224

Holman LE, 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. Sci Rep. https://doi.org/10.1038/s41598-019-47899-7

Intergovernmental Oceanographic Commission of UNESCO (2024) OBIS: Ocean Biodiversity Information System. obis.org

Invasive Species Specialist Group ISSG (2015) The Global Invasive Species Database. Version 2015.1. https://www.iucngisd.org/gisd/

Ito M, Onishi T, Dick MH (2015) Cribrilina mutabilis n. sp., an eelgrass-associated bryozoan (*Gymnolaemata*: *Cheilostomata*) with large variation in zooid morphology related to life history. Zool Sci 32(5):485–4975

Jerde CL, Chadderton WL, Mahon AR, Renshaw MA, Corush J, Budny ML, Mysorekar S, Lodge DM (2013) Detection of Asian carp DNA as part of a Great Lakes basin-wide surveillance program. Can J Fish Aquat Sci 70(4):522–526. https://doi.org/10.1139/cjfas-2012-0478

Jeunen GJ, Knapp M, Spencer HG, Taylor HR, Lamare MD, Stat M, Bunce M, Gemmell NJ (2019) Species-level biodiversity assessment using marine environmental DNA metabarcoding requires protocol optimization and standardization. Ecol Evol 9(3):1323–1335. https://doi.org/10.1002/ece3.4843

Kassambara A (2023) Ggpubr: "ggplot2" Based Publication Ready Plots (R package version 0.6.0). https://rpkgs.datanovia.com/ggpubr/

Katsanevakis S, Olenin S, Puntila-Dodd R, Rilov G, Stæhr PAU, Teixeira H, Tsirintanis K, Birchenough SNR, Jakobsen HH, Knudsen SW, Lanzén A, Mazaris AD, Piraino S, Tidbury HJ (2023) Marine invasive alien species in Europe: 9 years after the IAS Regulation. Front Marine Sci. https://doi.org/10.3389/fmars.2023.1271755

Kissling WD, Ahumada JA, Bowser A, Fernandez M, Fernández N, García EA, Guralnick RP, Isaac NJB, Kelling S, Los W, McRae L, Mihoub J-B, Obst M, Santamaria M, Skidmore AK, Williams KJ, Agosti D, Amariles D, Arvanitidis C, Bastin L, De Leo F, Egloff W, Elith J, Hobern D, Martin D, Pereira HM, Pesole G, Peterseil J, Saarenmaa H, Schigel D, Schmeller DS, Segata N, Turak E, Uhlir PF, Wee B, Hardisty AR (2018) Building essential biodiversity variables (EBVs) of species distribution and abundance at a global scale. Biol Rev 93:600–625. https://doi.org/10.1111/brv.12359

Knowlton N, Brainard RE, Fisher R, Moews M, Plaisance L, Caley MJ (2010) Coral reef biodiversity. Life World's. https://doi.org/10.1002/9781444325508.ch4

Kuhlenkamp R, Kind B (2018) Introduction of Non-indigenous Species. In: Salomon M, Markus T (eds) Handbook on Marine Environment Protection. Springer International Publishing, Cham, pp 487–516. https://doi.org/10.1007/978-3-319-60156-4_25

Lacoursière-Roussel A, Howland K, Normandeau E, Grey EK, Archambault P, Deiner K, Lodge DM, Hernandez C, Leduc N, Bernatchez L (2018) EDNA metabarcoding as a new surveillance approach for coastal Arctic biodiversity. Int J Bus Innov Res 17(3):7763–7777. https://doi.org/10.1002/ece3.4213

Lawson Handley L (2015) How will the "molecular revolution" contribute to biological recording? Biol J Lin Soc 115(3):750–766. https://doi.org/10.1111/bij.12516

Lee II H, Reusser D (2012) Atlas of Non-indigenous Marine and Estuarine Species in the North Pacific.

Lehtiniemi M, Ojaveer H, David M, Galil B, Gollasch S, McKenzie C, Minchin D, Occhipinti-Ambrogi A, Olenin S, Pederson J (2015) Dose of truth-monitoring marine non-indigenous species to serve legislative requirements. Mar Policy 54:26–35. https://doi.org/10.1016/j.marpol.2014.12.015

Leray M, Knowlton N (2015) DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. Proc Nat Acad Sci 112(7):2076–2081

Leray M, Knowlton N (2016) Censusing marine eukaryotic diversity in the twenty-first century. Philosophical Trans Royal Soc B Biol Sci. https://doi.org/10.1098/rstb.2015.0331

Leray M, Yang JY, Meyer CP, Mills SC, Agudelo N, Ranwez V, Boehm JT, Machida RJ (2013) A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. Front Zool 10(1):1–14. https://doi.org/10.1186/1742-9994-10-34/FIGURES/5

Leray M, Ho SL, Lin IJ, Machida RJ (2018) MIDORI server: a webserver for taxonomic assignment of unknown metazoan mitochondrial-encoded sequences using a curated database. Bioinformatics 34(21):3753–3754. https://doi.org/10.1093/BIOINFORMATICS/BTY454

Leray M, Knowlton N, Machida RJ (2022) MIDORI2: a collection of quality controlled, preformatted, and regularly updated reference databases for taxonomic assignment of eukaryotic mitochondrial sequences. Environ DNA 4(4):894–907. https://doi.org/10.1002/EDN3.303

Luypaert T, Hagan JG, McCarthy ML, Poti M (2020) Status of marine biodiversity in the Anthropocene. In: Jungblut S, Liebich V, Bode-Dalby M (eds) YOUMARES 9—The Oceans: Our Research, Our Future: Proceedings of the 2018 conference for Young Marine Researcher in Oldenburg, Germany. Springer International Publishing, Cham, pp 57–82. https://doi.org/10.1007/978-3-030-20389-4_4

Mahé F, Rognes T, Quince C, de Vargas C, Dunthorn M (2014) Swarm: robust and fast clustering method for amplicon-based studies. PeerJ 2014(1):e593. https://doi.org/10.7717/PEERJ.593/SUPP-8

Mahé F, Rognes T, Quince C, de Vargas C, Dunthorn M (2015) Swarmv2: highly-scalable and high-resolution amplicon clustering. PeerJ 2015(12):e1420. https://doi.org/10.7717/PEERJ.1420

Mahon AR, Jerde CL, Galaska M, Bergner JL, Chadderton WL, Lodge DM, Hunter ME, Nico LG (2013) Validation of eDNA surveillance sensitivity for detection of Asian carps in controlled and field experiments. PLoS ONE. https://doi.org/10.1371/journal.pone.0058316

Malani NV (2021) HiReadsProcessor: functions to process LM-PCR reads from 454/Illumina data.

Martaeng R, Obst M, Kuklinski P (2023) Phylogeographic study using autonomous reef monitoring structures indicates fast range expansion of the invasive bryozoan Juxtacribrilina mutabilis. Hydrobiologia 850(19):4115–4126. https://doi.org/10.1007/s10750-023-05184-9

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. http://www-huber.embl.de/users/an-

Massicotte P, South A (2024) Rnaturalearth: World map data from Natural Earth (R package version 1.0.1.9000). https://github.com/ropensci/rnaturalearth

Mazaris AD, Kallimanis A, Gissi E, Pipitone C, Danovaro R, Claudet J, Rilov G, Badalamenti F, Stelzenmüller V, Thiault L, Benedetti-Cecchi L, Goriup P, Katsanevakis S, Fraschetti S (2019) Threats to marine biodiversity in European protected areas. Sci Total Environ 677:418–426. https://doi.org/10.1016/j.scitotenv.2019.04.333

McInnes JC, Jarman SN, Lea MA, Raymond B, Deagle BE, Phillips RA, Catry P, Stanworth A, Weimerskirch H, Kusch A, Gras M, Cherel Y, Maschette D, Alderman R (2017) DNA metabarcoding as a marine conservation and management tool: a circumpolar examination of fishery discards in the diet of threatened albatrosses. Front Marine Sci. https://doi.org/10.3389/fmars.2017.00277

McMurdie PJ, Holmes S (2013) Phyloseq: an r package for reproducible interactive analysis and graphics of microbiome census data. PLoS ONE. https://doi.org/10.1371/journal.pone.0061217

Molnar JL, Gamboa RL, Revenga C, Spalding MD (2008) Assessing the global threat of invasive species to marine biodiversity. Front Ecol Environ 6(9):485–492. https://doi.org/10.1890/070064

Morgan M, Anders S, Lawrence M, Aboyoun P, Pagès H, Gentleman R (2009) ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. Bioinformatics 25:2607–2608. https://doi.org/10.1093/bioinformatics/btp450

EMorienLWParfrey2018SILVA v128 and v132 dada2 formatted 18s "train sets"Zenodo10.5281/zenodo.1447330Morien E, Parfrey LW (2018) SILVA v128 and v132 dada2 formatted 18s "train sets." Zenodo. https://doi.org/10.5281/zenodo.1447330

Neuwirth E (2014) RColorBrewer: ColorBrewer Palettes (R package version 1.1.3). https://CRAN.R-project.org/package=RColorBrewer

O'Hara CC, Frazier M, Halpern BS (2021) At-risk marine biodiversity faces extensive, expanding, and intensifying human impacts. Science 372(6537):84–87. https://doi.org/10.1126/science.abe6731

Obst M, Exter K, Allcock AL, Arvanitidis C, Axberg A, Bustamante M, Cancio I, Carreira-Flores D, Chatzinikolaou E, Chatzigeorgiou G, Chrismas N, Clark MS, Comtet T, Dailianis T, Davies N, Deneudt K, de Cerio OD, Fortič A, Gerovasileiou V, Pavloudi C (2020) A marine biodiversity observation network for genetic monitoring of hard-bottom communities (ARMS-MBON). Front Marine Sci. https://doi.org/10.3389/fmars.2020.572680

Occhipinti-Ambrogi A (2021) Biopollution by invasive marine non-indigenous species: A review of potential adverse ecological effects in a changing climate. Int J Environ Res Public. https://doi.org/10.3390/ijerph18084268

Othman N, Munian K, Haris H, Ramli FF, Sariyati NH, Najmuddin MF, Abdul-Latiff MAB (2023) A review on next-generation wildlife monitoring using environmental DNA (eDNA) detection and next-generation sequencing in malaysia. Sains Malaysiana 52(1):17–33

Pagad S, Genovesi P, Carnevali L, Schigel D (2018) Introducing the global register of introduced and invasive species. Sci Data 5:170202. https://doi.org/10.1038/sdata.2017.202

Pagès H, Aboyoun P, Gentleman R, DebRoy S (2020) Biostrings: Efficient manipulation of biological strings. https://bioconductor.org/packages/Biostrings

Palomares MLD, Pauly D (2024) SeaLifeBase. www.sealifebase.org

Patrício J, Little S, Mazik K, Papadopoulou KN, Smith CJ, Teixeira H, Hoffmann H, Uyarra MC, Solaun O, Zenetos A, Kaboglu G, Kryvenko O, Churilova T, Moncheva S, Bucas M, Borja A, Hoepffner N, Elliott M (2016) European marine biodiversity monitoring networks: strengths, weaknesses, opportunities and threats. Front Marine Sci. https://doi.org/10.3389/fmars.2016.00161

Pearman JK, Chust G, Aylagas E, Villarino E, Watson JR, Chenuil A, Borja A, Cahill AE, Carugati L, Danovaro R, David R, Irigoien X, Mendibil I, Moncheva S, Rodríguez-Ezpeleta N, Uyarra MC, Carvalho S (2020) Pan-regional marine benthic cryptobiome biodiversity patterns revealed by metabarcoding autonomous reef monitoring structures. Mol Ecol 29(24):4882–4897. https://doi.org/10.1111/mec.15692

Pearman JK, von Ammon U, Laroche O, Zaiko A, Wood SA, Zubia M, Planes S, Pochon X (2021) Metabarcoding as a tool to enhance marine surveillance of nonindigenous species in tropical harbors: a case study in Tahiti. Environ DNA 3(1):173–189. https://doi.org/10.1002/edn3.154

Plaisance L, Brainard R, Julian Caley M, Knowlton N (2011) Using DNA barcoding and standardized sampling to compare geographic and habitat differentiation of crustaceans: a Hawaiian Islands example. Diversity 3(4):581–591. https://doi.org/10.3390/d3040581

Poorter De M, Darby C, MacKay J (2008) Marine Menace. marine menace—alien invasive species in the marine environment, 1–32. http://www.issg.org/pdf/publications/marine_menace_english.pdf

Portas A, Quillien N, Culioli G, Briand JF (2022) Eukaryotic diversity of marine biofouling from coastal to offshore areas. Front Marine Sci. https://doi.org/10.3389/fmars.2022.971939

Porter TM, Hajibabaei M (2020) Putting COI metabarcoding in context: the utility of exact sequence variants (ESVs) in biodiversity analysis. Front Ecol Evol 8:248. https://doi.org/10.3389/FEVO.2020.00248/BIBTEX

Pyšek P, Hulme PE, Simberloff D, Bacher S, Blackburn TM, Carlton JT, Dawson W, Essl F, Foxcroft LC, Genovesi P, Jeschke JM, Kühn I, Liebhold AM, Mandrak NE, Meyerson LA, Pauchard A, Pergl J, Roy HE, Seebens H, Richardson DM (2020) Scientists' warning on invasive alien species. Biol Rev 95(6):1511–1534

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41(D1):D590–D596. https://doi.org/10.1093/NAR/GKS1219

R Core Team (2021) R: a language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org

Radashevsky VI, Malyar VV, Pankova VV, Sato-Okoshi W (2020) Disentangling invasions in the sea: molecular analysis of a global polychaete species complex (*Annelida*: *Spionidae*: *Pseudopolydora paucibranchiata*). Biol Invasions 22(11):3621–3644. https://doi.org/10.1007/s10530-020-02346-x

Ranwez V, Douzery EJP, Cambon C, Chantret N, Delsuc F (2018) MACSE v2: toolkit for the alignment of coding sequences accounting for frameshifts and stop codons. Mol Biol Evol 35(10):2582–2584. https://doi.org/10.1093/molbev/msy159

Ratnasingham S, Hebert PDN (2007) Bold: the barcode of life data syste. Mol Ecol Notes 7(3):355–364

Roman J, Palumbi SR (2004) A global invader at home: population structure of the green crab, *Carcinus maenas*. Europe Mol Ecol 13(10):2891–2898. https://doi.org/10.1111/j.1365-294X.2004.02255.x

Ruppert KM, Kline RJ, Rahman MS (2019) Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: a systematic review in methods, monitoring, and applications of global eDNA. Global Ecol Conserv. https://doi.org/10.1016/j.gecco.2019.e00547

Ryland JS, Bishop JDD, de Blauwe H, El Nagar A, Minchin D, Wood CA, Yunnie ALE (2011) Alien species of *Bugula* (*Bryozoa*) along the Atlantic coasts of Europe. Aquat Invasions 6(1):17–31. https://doi.org/10.3391/ai.2011.6.1.03

Santi I, Beluche O, Beraud M, Buttigieg PL, Casotti R, Cox CJ, Cunliffe M, Davies N, de Cerio OD, Exter K, Kervella AE, Kotoulas G, Lagaisse R, Laroquette A, Louro B, Not F, Obst M, Pavloudi C, Poulain J, Præbel K, Vanaverbeke J, Pade N (2023) European marine omics biodiversity observation network: a strategic outline for the implementation of omics approaches in ocean observation. Front Mar Sci 10:1118120. https://doi.org/10.3389/fmars.2023.1118120

Shenkar N, Loya Y (2008) The solitary ascidian *Herdmania momus*: native (Red Sea) versus non-indigenous (Mediterranean) populations. Biol Invasions 10(8):1431–1439. https://doi.org/10.1007/s10530-008-9217-2

Siegenthaler A, Wangensteen OS, Soto AZ, Benvenuto C, Corrigan L, Mariani S (2019) Metabarcoding of shrimp stomach content: harnessing a natural sampler for fish biodiversity monitoring. Mol Ecol Resour 19(1):206–220. https://doi.org/10.1111/1755-0998.12956

Stefanni S, Stanković D, Borme D, de Olazabal A, Juretić T, Pallavicini A, Tirelli V (2018) Multi-marker metabarcoding approach to study mesozooplankton at basin scale. Sci Rep. https://doi.org/10.1038/s41598-018-30157-7

Takahara T, Minamoto T, Doi H (2013) Using environmental DNA to estimate the distribution of an invasive fish species in ponds. PLoS ONE. https://doi.org/10.1371/journal.pone.0056584

Tanhua T, Pouliquen S, Hausman J, O'Brien KM, Bricher P, de Bruin T, Buck JJ, Burger EF, Carval T, Casey KS, Diggs S, Giorgetti A, Glaves H, Harscoat V, Kinkade D, Muelbert JH, Novellino A, Pfeil BG, Pulsifer P, Zhao Z (2019) Ocean FAIR data services. Front Marine Sci. https://doi.org/10.3389/fmars.2019.00440

Tenopir C, Rice NM, Allard S, Baird L, Borycz J, Christian L, Grant B, Olendorf R, Sandusky RJ (2020) Data sharing, management, use, and reuse: practices and perceptions of scientists worldwide. PLoS ONE. https://doi.org/10.1371/journal.pone.0229003

Thomas AC, Deagle BE, Eveson JP, Harsch CH, Trites AW (2016) Quantitative DNA metabarcoding: improved estimates of species proportional biomass using correction factors derived from control material. Mol Ecol Resour 16(3):714–726. https://doi.org/10.1111/1755-0998.12490

Thomasdotter A, Shum P, Mugnai F, Vingiani M, Dubut V, Marschal F, Abbiati M, Chenuil A, Costantini F (2023) Spineless and overlooked: DNA metabarcoding of autonomous reef monitoring structures reveals intra- and interspecific genetic diversity in Mediterranean invertebrates. Mol Ecol Resour 23(7):1689–1705. https://doi.org/10.1111/1755-0998.13836

Thomsen PF, Willerslev E (2015) Environmental DNA—an emerging tool in conservation for monitoring past and present biodiversity. Biol Conserv. https://doi.org/10.1016/j.biocon.2014.11.019

Trebitz AS, Hoffman JC, Darling JA, Pilgrim EM, Kelly JR, Brown EA, Chadderton WL, Egan SP, Grey EK, Hashsham SA, Klymus KE, Mahon AR, Ram JL, Schultz MT, Stepien CA, Schardt JC (2017) Early detection monitoring for aquatic non-indigenous species: optimizing surveillance, incorporating advanced technologies, and identifying research needs. J Environ Manag. https://doi.org/10.1016/j.jenvman.2017.07.045

Tréguier A, Paillisson JM, Dejean T, Valentini A, Schlaepfer MA, Roussel JM (2014) Environmental DNA surveillance for invertebrate species: advantages and technical limitations to detect invasive crayfish Procambarus clarkii in freshwater ponds. J Appl Ecol 51(4):871–879. https://doi.org/10.1111/1365-2664.12262

Trujillo-González A, Li T, Potts J, Nicol S, Allain V, Godwin SC, Vourey E, Portal A, Kumasi B, Usu T, Rodrigo A, Gleeson D (2022) Can stomach content and microbiomes of tuna provide near real-time detection of ecosystem composition in the pacific ocean? Front Marine Sci. https://doi.org/10.3389/fmars.2022.811532

Tsiamis K, Azzurro E, Bariche M, Çinar ME, Crocetta F, De Clerck O, Galil B, Gómez F, Hoffman R, Jensen KR, Kamburska L, Langeneck J, Langer MR, Levitt-Barmats

Y, Lezzi M, Marchini A, Occhipinti-Ambrogi A, Ojaveer H, Piraino S, Cardoso AC (2020) Prioritizing marine invasive alien species in the European Union through horizon scanning. Aquatic Conserv Marine Freshwater Ecosyst 30(4):794–845

Verones F, Gjedde P, Koslowski M, Woods JS, Lonka R, Stadler K (2023) MarINvaders: a web toolkit of marine species for use in environmental assessments. Ecosphere 14(11):e4697. https://doi.org/10.1002/ecs2.4697

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

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 73(16):5261–5267. https://doi.org/10.1128/AEM.00062-07

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

West K, Travers MJ, Stat M, Harvey ES, Richards ZT, DiBattista JD, Newman SJ, Harry A, Skepper CL, Heydenrych M, Bunce M (2021) Large-scale eDNA metabarcoding survey reveals marine biogeographic break and transitions over tropical north-western Australia. Divers Distrib 27(10):1942–1957. https://doi.org/10.1111/ddi.13228

Westfall KM, Therriault TW, Abbott CL (2020) A new approach to molecular biosurveillance of invasive species using DNA metabarcoding. Glob Change Biol 26(2):1012–1022. https://doi.org/10.1111/gcb.14886

Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen T, Miller E, Bache S, Müller K, Ooms J, Robinson D, Seidel D, Spinu V, Yutani H (2019) Welcome to the tidyverse. J Open Source Softw 4(43):1686

Wickham H, Hester J, Chang W, Bryan J (2022) Devtools: Tools to Make Developing R Packages Easier

Wickham H, François R, Henry L, Müller K (2023) Dplyr: A Grammar of Data Manipulation. R package version 1.1.2. https://cran.r-project.org/package=dplyr

Wickham H, Pedersen T, Seidel D (2023) Scales: Scale Functions for Visualization (R package version 1.3.0). https://scales.r-lib.org

Wickham H, Vaughan D, Girlich M (2023) Tidyr: Tidy Messy Data. https://cran.r-project.org/package=tidyr

Wickham H, Vaughan D, Girlich M (2024) Tidyr: Tidy Messy Data (R package version 1.3.1). https://github.com/tidyverse/tidyr

Wickham H (2016) Ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wickham H (2022) Stringr: Simple, Consistent Wrappers for Common String Operations. https://cran.r-project.org/package=stringr

Wickham H (2023) Forcats: Tools for Working with Categorical Variables (Factors) (R package version 1.0.0). https://github.com/tidyverse/forcats

Wu Y, Colborne SF, Charron MR, Heath DD (2023) Development and validation of targeted environmental DNA (eDNA) metabarcoding for early detection of 69 invasive fishes and aquatic invertebrates. Environ DNA 5(1):73–84. https://doi.org/10.1002/edn3.359

Yu DW, Ji Y, Emerson BC, Wang X, Ye C, Yang C, Ding Z (2012) Biodiversity soup: Metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. Methods Ecol Evol 3(4):613–623. https://doi.org/10.1111/j.2041-210X.2012.00198.x

Yuan D, Ahamed A, Burgin J, Cummins C, Devraj R, Gueye K, Gupta D, Gupta V, Haseeb M, Ihsan M, Ivanov E, Jayathilaka S, Kadhirvelu VB, Kumar M, Lathi A, Leinonen R, McKinnon J, Meszaros L, O'Cathail C, Cochrane G (2024) The European nucleotide archive in 2023. Nucleic Acids Res 52(D1):D92–D97. https://doi.org/10.1093/nar/gkad1067

Zarcero J, Antich A, Rius M, Wangensteen OS, Turon X (2024) A new sampling device for metabarcoding surveillance of port communities and detection of non-indigenous species. Iscience. https://doi.org/10.1016/j.isci.2023.108588

Zhan A, Xiong W, He S, MacIsaac HJ (2014) Influence of artifact removal on rare species recovery in natural complex communities using high-throughput sequencing. PLoS ONE. https://doi.org/10.1371/journal.pone.0096928