Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Integrative Taxonomy Improves Biodiversity Characterization of Vulnerable Marine Ecosystems at Flemish Cap

Marina Parrondo
parrondomarina@proton.me
Instituto de Investigaciones Marinas
Bruno Almón
Spanish Institute of Oceanography
Nair Vilas-Arrondo
Instituto de Investigaciones Marinas
Diana Casado-Albarrán
Instituto de Investigaciones Marinas
Francisco González-Carrión
Instituto de Investigaciones Marinas
Celso Domingos
University of Porto
Andreu Santín
University of Porto
Celeste L. Hinojo
Instituto de Investigaciones Marinas
José Miguel Casas-Sánchez
Spanish Institute of Oceanography
Fran Saborido-Rey
Instituto de Investigaciones Marinas
Laura Casas
Instituto de Investigaciones Marinas

Article

Keywords: Barcoding, Benthic biodiversity, Cold-water corals, Deep-sea, Northwest Atlantic, Sponge aggregations.

Posted Date: July 14th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9926518/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Recognized as biodiversity hotspots and by their high susceptibility to human activities, Vulnerable Marine Ecosystems (VMEs) are increasingly targeted for protection from fishing practices, particularly bottom trawling, which may impact benthic communities. Safeguarding these ecosystems requires non-destructive tools to describe and map seabed biodiversity at a faster pace than current imaged-based technologies. Environmental DNA (eDNA) is a promising approach, requiring only water sampling. However, its effectiveness relies on public DNA databases, which remain incomplete and error-prone, particularly for invertebrates in understudied habitats. One such area is the Flemish Cap, an isolated seamount with a long history of exploitation and home to several VMEs species indicators, yet still poorly documented. To address this gap, a 2022 survey collected and identified marine invertebrates through classical taxonomy and DNA barcoding (COI and 16S), providing essential reference data to support non-invasive conservation strategies. Morphological analysis identified 91 morphospecies and 63 low-resolution categories among 434 specimens. DNA barcoding revealed that although over 85% of sequences matched public database entries at 97% identity, only 30% yielded consistent assignments across markers and repositories. Further taxonomic revision improved the classification of 46.6% of specimens, identifying 49 additional species, thereby expanding the reference library to 127 species. Persistent challenges remain for Porifera and Cnidaria, reflecting database gaps and amplification biases. Our results demonstrate that integrative taxonomy markedly improves assignment success and provides essential baseline data for deep-sea biodiversity assessments. By strengthening molecular reference resources for VME indicator taxa, this study directly supports the development of robust, non-invasive ecological indicators for long-term monitoring and conservation of vulnerable deep-sea habitats.

INTRODUCTION

The deep sea, usually defined as waters below 200 m depth, is one of the largest and most diverse ecosystems on Earth[1,2]. It harbours unique and fragile communities, many of which form structurally complex habitats that support high biodiversity and provide essential ecosystem services to humans[3,4]. Among these systems, Vulnerable Marine Ecosystems (VMEs) are particularly relevant. They include habitats dominated by species with low resistance to disturbance and limited recovery capacity[5]. VMEs are often defined by their structural complexity, and key ecological functions that sustain rich biodiversity. Examples of VMEs indicator species (VMEIs) include cold-water corals[6,7] sponge-dominated communities and structural biogenic habitats[8,9]; as well as endemic or rare types of hydrothermal vent and cold seep communities[10,11]. Because of their fragility and critical role in maintaining ecosystem integrity, VMEs are a priority for conservation and management of the deep-sea fisheries, especially in the Areas Beyond National Jurisdiction (ABNJ) regulated by Regional Fisheries Management Organizations (RFMOs)[12].

Epibenthic megafauna are key components of VMEs. They contribute substantially to benthic biomass and link benthic and pelagic ecosystems[13]. These organisms not only increase habitat complexity through their physical structures[14,15], but also play a role in altering small-scale habitat structures and biochemical conditions[16]. They also support higher trophic levels such as fish and marine mammals, and serve as indicators of fishing impacts[17].

Despite their ecological importance, basic biological and ecological information remains limited for most VMEIs. This gap includes limited knowledge on species ranges, population structure, genetic diversity, dispersal capacity and demographic parameters, all of which hinders the implementation of effective conservation measures by RFMOs[18,19]. Currently, the majority of the information of VMEs comes from fisheries-dependent sources, such as trawl and longline bycatch, which have an impact on these habitats. This highlights the need to collect non-invasive data for VME identification and monitoring[18]. Current methods include kernel density estimates of VMEIs from trawl surveys[20] and image-based multi-criteria assessments[21,22]. However, both approaches have limitations in resolution, scalability or impact.

The use of molecular tools has significantly improved the ability of scientists and managers to understand, conserve and sustainably manage marine resources, contributing to the preservation of marine ecosystems[23]. One of the molecular techniques that has recently attracted attention is environmental DNA (eDNA) metabarcoding, which targets DNA extracted directly from environmental samples (such as soil, water or air) without isolating organisms[24]. eDNA represents a complex mixture of genomic DNA from many taxa and degradation states offering a fast, non-invasive, efficient, and cost-effective approach that involves only non-destructive methods while avoiding significant damage to the target species or its habitats. It enables simultaneous biodiversity assessment across a wide range of taxa[25–27] and even has demonstrated capacity to estimate abundance and biomass of marine organisms[28,29], although such inferences remain context dependent. Therefore, eDNA-based techniques show strong potential for the study and management of marine benthic habitats and VMEs. However, its effectiveness depends on the completeness and accuracy of genetic information in public DNA databases, which remain limited, potentially leading to assignment errors and detection failures[30–32]. Current molecular information on marine organisms remain incomplete and taxonomically biased. In particular, marine invertebrates are frequently overlooked, in spite of their functional importance and exceptional diversity[33,34]. This has resulted in a major lack of genetic and genomic baseline data[35,36].

In the Northwest Atlantic Fisheries Organization (NAFO) regulatory area, several VMEs have been identified in both fished[37] and unfished regions[38]. Identification has largely relied on kernel density estimates of VME indicators derived from EU groundfish research surveys[39]. These estimates have informed the closure of 18 areas to bottom fishing, most of which are located in Flemish Cap (NAFO 3M Division), to prevent impacts such as by-catch and discarding of corals and sponges. This area is a dome-shaped, deep-water seamount, separated from the Newfoundland shelf by the > 1100 m deep Flemish Pass, which limits faunal exchange with the Grand Banks. Since 1988, an annual bottom trawl research survey has been conducted in Flemish Cap under the European Union Data Collection Framework[40]. This survey primarily assesses stock biomass and abundance indexes of commercial species, but also records all species captured, including epibenthic invertebrates. Due to the low selectivity of trawl gear, catches encompass a broad range of faunal groups, providing a comprehensive overview of the area's biodiversity. While these surveys supply valuable data, they rely on invasive methods and involve the sacrifice of VME indicator taxa.

Recently, NAFO has been exploring the potential use of non-invasive sampling techniques to monitor VMEs[41,42], as it remains one of the few RFMOs that assess VMEs primarily through invasive bottom trawl research surveys. Incorporating non-invasive methods into VME monitoring programs could strengthen NAFO's research capabilities, provide a model for VME studies in other regions, and enable large-scale, sustainable ecosystem observation.

In this context, the objective of the present study was to develop a taxonomically comprehensive barcode library of deep-sea invertebrates inhabiting Flemish Cap. Accurate species identification using short DNA fragments is essential to develop non-destructive tools that can reliably assess benthic communities in isolated areas, establish biodiversity baselines, monitor long-term changes, and protect these ecosystem engineers from the impacts of harmful human activities.

RESULTS

Morphological identification

A total of 434 epibenthic invertebrates were classified morphologically into 154 unique taxonomic categories (Fig. 1). These comprised 91 morphospecies (59%), while the remaining 63 taxa (41%) could not be identified to the species level. Of these, 37 were assigned at genus level (24%) and 26 to broader taxonomic categories (17%). These 63 taxa were designated as Low Taxonomic Resolution (LTR), following Keck et al., (2023)[43].

Most morphospecies had at least one representative COI sequence available in public databases. Specifically, 73.62% were represented in BOLD and 76.92% in GenBank, with 68.13% represented in both repositories. For 16S, 53.84% of morphospecies had a reference sequence in GenBank. A substantial proportion of these references were singletons (i.e., represented by only one sequence), accounting for 3.30% of species in BOLD, 9.89% in GenBank for COI, and 24.17% in GenBank for 16S (Fig. 2). Overall, 17.58% of morphospecies lacked a COI reference sequence, and 46.15% lacked a 16S reference sequence in public databases.

Sequence representation was uneven across taxonomic groups (Fig. 3). A clear taxonomic bias was evident in public sequence databases, with arthropods showing the highest representation for both markers (BOLD: 75%; GenBank COI: 78.12%; GB_16S: 46.87%), followed by mollusks (BOLD: 45.83%; GenBank COI: 50%; GenBank 16S: 41.66%) and echinoderms (BOLD: 48.38%; GenBank COI: 41.93%; GenBank 16S: 32.25%) (Fig. 3). Specimens designated as LTR during morphological identification accounted for 40.91% of all unique taxonomic categories and comprised more than 50% of samples within the phyla Annelida (75%), Chaetognatha (100%), Chordata (100%), Cnidaria (50%) and Porifera (60.71%) (Fig. 4). The lack of taxonomic identification prevented their inclusion in searches against public genomic databases. Morphospecies assignments to BINs revealed a largely balanced distribution. Concordant and discordant BINs each accounted for 29.67% of the cases, multiple BINs comprised 10.99%, and unclassified or NA values represented the remaining 29.67%.

DNA barcoding for molecular identification

DNA was successfully extracted from all 434 individuals. Universal primers effectively amplified most taxonomic groups but amplification success varied considerably between markers across groups (COI and 16S, Table 1, Fig. 4). For COI, a ~ 640 bp fragment of the barcode region was successfully amplified in 342 samples (79%). All sequences were functional and showed no insertions, deletions, or stop codons. For 16S, a ~ 580 bp fragment of the rRNA gene was obtained from 258 samples (59%).

Table 1
Summary of technical success in PCR amplification and sequencing for COI and 16S markers. The table shows the number of specimens (n) and their relative percentage according to amplification outcome combinations for both markers.

Amplification COI	Amplification 16S	N	Percentage %
Success	Success	214	49.3
Success	Failure	128	29.5
Failure	Success	44	10.1
Failure	Failure	48	11.1
		434	**100**

A comparison of the resulting sequences against BOLD and GenBank databases showed that 85.8% met the 97% pairwise identity threshold for at least one marker in either database. Only 30.57% were successfully assigned across both markers and databases simultaneously. The remaining 14.2% fell below the 97% threshold, representing potential new genetic records for deep-sea macroinvertebrates (Fig. 5).

Amplification and database assignment success varied markedly at the order level, with clear differences among phyla (Fig. 6). Arthropoda and Cnidaria were the most represented groups, with 67 and 113 specimens, respectively. Both achieved very high COI amplification success (> 98%) while 16S performance was markedly lower (64.2% and 35.4%, respectively. Within Arthropoda, Decapoda (n = 45) dominated, showing 97.8% COI and 73.3% 16S amplification. Smaller orders such as Lophogastrida (n = 9) and Scalpellomorpha (n = 5) exhibited 100% COI amplification but reduced 16S success (11.1% and 80%, respectively). In Cnidaria, Actiniaria (n = 38) and Scleractinia (n = 12) achieved 100% COI but 0% 16S amplification, whereas Scleralcyonacea (n = 35) showed high success for both markers (100% COI, 88.6% 16S). Database assignment success was generally high for COI in BOLD (100%) and GenBank (80–90%), while 16S matches were more variable (51–83%). Echinodermata displayed the opposite trend, with moderate COI (63.8%) but near-complete 16S

amplification (96.3%). This pattern was consistent across orders with Valvatida (n = 27) showing 55.6% COI vs 100% 16S, Paxillosida (n = 15) 53.3% COI vs 100% 16S, and Ophiurida (n = 8) 50% COI vs 100% 16S. Echinothurioida (n = 7) and Forcipulatida (n = 5) achieved 100% amplification for both markers. Database assignments exceeded 70% in both GenBank and BOLD. Mollusca displayed balanced performance across both markers (COI 95.6%, 16S 84.4%), with high assignment rates in all databases (> = 72%). The dominant orders, Oegopsida (n = 13), Neogastropoda (n = 10), and Octopoda (n = 10), showed COI amplification rates ≥ 90%, with the first two achieving 100%. For 16S, amplification ranged from 92.3% to 100%, indicating consistent marker performance across major molluscan lineages. Porifera showed 86.6% COI and 77.6% 16S amplification, but GenBank assignment for 16S was low (32.7%). Tetractinellida (n = 40) showed 90% COI and 95% 16S amplification, yet only 28.9% of 16S sequences matched GenBank references, while Polymastiida (n = 7) showed 100% COI but failed entirely for 16S. Annelida achieved 83.3% amplification for both markers, but only 40% of 16S sequences matched GenBank. Phyllodocida (n = 4) had 75% success for both markers, and Eunicida (n = 1) achieved 100% amplification and assignment. Chordata, Brachiopoda, and Chaetognatha all showed 100% COI amplification. However, 16S amplification failed in Chordata and Chaetognatha, while Brachiopoda achieved full amplification but with variable assignment success (0-100%). Within Chordata, Stolidobranchia (n = 3) and Aplousobranchia (n = 1) amplified COI (100%) but failed 16S (0%), a pattern also observed in Brachiopoda (Terebratulida, n = 3) and Chaetognatha (Aphragmophora, n = 1).

Integrative taxonomy

An integrative taxonomy approach combining expert morphological identification with molecular data was applied to all specimens, allowing comparison between the initial 154 morphologically defined taxonomic categories and their corresponding genetic identities. Of these, 43.5% showed full congruence between morphological and molecular identifications. In contrast, 45.6% displayed mismatches between the two approaches. A third group, comprising 10.9% of the specimens, did not reach the 97% similarity threshold for any marker in public databases and, therefore, could not be assigned to an existing reference sequence. The proportion of specimens requiring expert revision varied across phyla. All chordate specimens (100%) required additional evaluation, followed by Porifera (73.1%), Mollusca (37.8%), Annelida (33.3%), Echinodermata (22.5%), Arthropoda (20.9%), and Cnidaria (17.7%). No specimens belonging to Brachiopoda or Chaetognatha required revision, although these phyla were represented by small sample sizes.

The integrative taxonomic approach resulted in variable changes in taxonomic resolution across phyla, as illustrated in matrices comparing morphological identification (rows) with post-integration assignments (columns) (Fig. 7A-B). Overall, 180 out of 386 specimens (46.6%) improved their taxonomic resolution through integration, 246 specimens (63.7%) retained their initial resolution, and 10 specimens (2.6%) were assigned to a higher level (Fig. 7A). Porifera exhibited the highest improvement among well-represented phyla, with 53.7% of specimens (n = 36/67) reaching a lower taxonomic level (median taxonomic improvement of 1 level). Mollusca, Cnidaria, Echinodermata, and Arthropoda exhibited more modest improvement rates of 37.8%, 36.3%, 23.8%, and 11.9%, respectively, with median improvements of 0 levels (Fig. 7B). Phyla represented by few specimens showed variable outcomes. Chaetognatha (n = 1) and Chordata (n = 4) achieved the highest proportional improvement for all specimens with a median improvement of 5 and 2 levels, respectively. Annelida (n = 6) showed a median improvement of 1 level in 66.7% of the specimens (Fig. 7B) whereas Brachiopoda (n = 3) showed no changes in resolution (Fig. 7).

Integrative taxonomic refinement resulted in the identification of 42 additional morphospecies (49 when including open taxonomy), increasing the number from 91 initially identified morphospecies to 120 (127 species considering open taxonomy). Our study also revealed significant data gaps in public sequence databases. After initial morphological assignment, 70 unique taxonomic categories lacked reference sequences or could not be retrieved due to classification as LTR. Following integrative refinement, this number increased to 85 unique taxonomic categories, which now have a representative sequence in public genomic databases.

DISCUSSION

The present study aimed to enhance DNA reference databases for marine invertebrates from a poorly characterized deep-sea habitat by employing an integrative taxonomic approach to build a robust reference library. By refining morphospecies assignments and resolving cryptic diversity, this approach addresses the limitations of relying solely on DNA barcoding, which remains insufficient for some groups due to marker variability, incomplete reference libraries, and persistent taxonomic gaps[44–46]. Our findings underscore the importance of developing local DNA sequence libraries to complement incomplete or potentially inaccurate reference databases[47–49]. Strengthening such resources is critical for the development of non-destructive monitoring tools and for accurate biodiversity assessments of epibenthic macroinvertebrates in the Flemish Cap region. This need is particularly pressing because Flemish Cap hosts a high concentration of VMEs, including coral gardens and sponge grounds, which play key ecological roles yet remain highly sensitive to disturbance[50–53]. Increasing anthropogenic pressures, such as bottom-contact fishing, deep sea mining, and climate change, further underscore the importance of establishing robust biodiversity baselines to guide conservation and sustainable management.

Despite recent progress, marine invertebrates remain markedly underrepresented in reference databases, and many marine shelf, deep-water, and polar areas continue to be insufficiently sampled[54,55]. Over the last decade, barcode libraries have expanded through both targeted projects and international initiatives focusing on specific groups, including Annelida[56], Mollusca[48,57,58], Arthropoda[59] and Porifera[60]. Nevertheless, substantial gaps remain, with molecular coverage sometimes falling below 30% for several deep-sea and understudied taxa[30,61]. These gaps stem from multiple, well-documented factors, including insufficient barcode coverage[30,62], uneven taxonomic representation across phyla and regions[61], and the presence of misidentified, low-quality, or non-target sequences in public repositories[63]. Collectively, these limitations hinder accurate species delimitation and reduce the effectiveness of molecular biodiversity assessments.

Consistent with these broader database limitations, COI reference sequences were available for just over 70% of the morphospecies analysed in our study, whereas only 54% had representative 16S sequences. Importantly, even when reference data were present, our integrative analysis revealed notable discrepancies, with 45.6% of the specimens showing mismatches between morphological and molecular identifications. These inconsistencies likely reflect

cryptic diversity and the known prevalence of morphologically indistinguishable lineages in sponges and deep-sea corals, as well as potential misidentifications arising from the plasticity of morphological traits in these groups[64–66]. In addition, they may reflect the uneven geographic and taxonomic representation of deep-sea taxa, as well as the incomplete coverage and variable quality of reference sequences available in public genetic databases[67–69]. Additionally, more than 10% of specimens lacked any reference sequence above the similarity threshold[70], further underscoring the extent to which many deep-sea lineages remain molecularly undersampled.

Amplification success also varied markedly across phyla. Porifera, polychaetes, and ascidians exhibited reduced COI amplification, consistent with previously reported success rates ranging from ~ 20% in ascidians and sponges to ~ 50% in annelids[71–73]. Echinoderms showed over 35% COI amplification failure, confirming earlier observations of poor COI performance in this group[74,75]. In contrast, 16S primers performed relatively well, supporting previous work highlighting the advantages of a multigene approach for improving identification success[48,76]. Cnidarians displayed high COI amplification success and database matches, yet low mitochondrial divergence and pronounced morphological plasticity limited species-level resolution, consistent with previous studies[77–79]. A similar pattern was observed in Porifera, and, to a lesser extent, in Echinodermata. These findings align with earlier studies reporting that slow mitochondrial evolution and the presence of pseudogenes can restrict interspecific divergence in these groups, thereby reducing the species-level discriminatory power of COI[80].

Additional complications arose from the recovery of non-target sequences in filter-feeding taxa such as corals, sponges, or ascidians, where DNA from holobionts (Pseudomonadota and Proteobacteria) and co-inhabiting organisms was occasionally amplified. Although relatively uncommon, these events were more frequent with 16S and are consistent with challenges in barcoding filter feeders documented in previous studies[73].

The integrative approach also revealed substantial differences among phyla in the proportion of specimens requiring expert revision. These patterns likely reflect variation in morphological complexity, taxonomic difficulty, and the prevalence of poorly resolved or highly diverse groups. The substantial revision required in Chordata and Porifera is consistent with well-known taxonomic challenges in these phyla, where morphology-based identification is often difficult. Cryptic diversity, phenotypic plasticity, unresolved systematics, and the limited availability of reliable reference sequences complicate the resolution of species boundaries[81–83]. Similarly, the considerable revision needed for Mollusca and Annelida aligns with long-recognized taxonomic difficulties in these groups, as well as with gaps and ambiguities in DNA reference libraries[69,84–86]. Despite these challenges, the integrative approach markedly improved taxonomic resolution, particularly in Porifera and in phyla with small sample sizes, where individual specimens benefited substantially from the combined evidence. In contrast, groups such as Cnidaria, Echinodermata, Mollusca, and Arthropoda showed more limited gains, suggesting that morphological identifications were already close to their achievable resolution or that molecular markers alone were insufficient for further refinement due to incomplete databases. Overall, 90.5% of LTR cases were successfully resolved, resulting in the identification of 42 additional morphospecies (49 including open taxonomy) and expanding the reference library from 91 to 120 species (127 including open taxonomy).

Integrative taxonomy, combining molecular, morphological, ecological, and biogeographical data, provides a more robust framework to resolve species boundaries and refine morphospecies assignments[87,88]. However, this approach remains heavily dependent on taxonomic expertise, which is increasingly scarce, and further constrained by the decline in taxonomic positions since the 1990s[89–91]. As a result, deep-sea biodiversity remains largely uncharacterized and molecular tools alone are insufficient to resolve these challenges. By providing a comprehensive characterization of epibenthic macroinvertebrates from Flemish Cap, this study contributes to filling these critical gaps. The reference library developed here enhances the accuracy of biodiversity assessments, supports the development of non-destructive monitoring tools, and strengthens the scientific basis for the sustainable management and protection of VME indicator species. The future expansion of this library, together with the generalization of the environmental DNA approach, will enable progress towards more sensitive, comparable and operational tools for assessing ecological status in deep-sea waters, thus contributing to the establishment of early-warning indicators for ecosystem degradation.

CONCLUSIONS

Integrative taxonomy proved essential for resolving species-level diversity on Flemish Cap, where incomplete reference databases, cryptic lineages, and morphologically challenging taxa hinder accurate identification. The combination of molecular, morphological, and ecological data substantially improved the characterization of the epibenthic macroinvertebrate community and expanded the local reference library, providing a more reliable baseline for biodiversity assessments in this remote deep-sea region. These advances not only enhance the accuracy of biodiversity species assessments in Flemish Cap but also support the development of non-destructive monitoring tools tailored to its vulnerable marine ecosystems. As human pressures intensify across deep-sea habitats, robust taxonomic and molecular resources such as those developed here are essential to guide conservation, sustainable management, and the protection of ecologically critical species and communities. The implementation of this framework would enhance NAFO's and other RFMOs ability to carry out adaptive, ecosystem-based management, ultimately contributing to the long-term conservation of vulnerable deep-sea habitats.

MATERIALS AND METHODS

Study area, sample collection and morphological identification

The data and samples used in the present study were collected during the 2022 European Union Flemish Cap bottom-trawl survey aboard the research vessel (RV) *Vizconde de Eza* between July and August 2022. Sampling followed a stratified random design using a Lofoten bottom trawl gear. A total of 183 daytime trawls were conducted, each lasting 30 minutes, covering the entire bank down to depths of 1460 meters (Fig. 8).

All epibenthic invertebrates retained by the trawl were sorted and identified on board to the lowest possible taxonomic level. For each taxon, both the number of individuals and their wet weight were recorded, except for colonial organisms, which were computed as a single organism, with the weight representing the entire colony. A total of 434 individuals were collected for this study, labelled and stored at -20°C until further processing.

Upon arrival at the laboratory, specimens were preserved in 95% ethanol for subsequent genomic analysis. A taxonomic revision of the on-board sample classification was conducted to confirm identifications and refine taxonomic assignments. The resulting taxa checklist was then cross-referenced against public genomic databases to confirm the availability of representative sequences for the regions targeted in subsequent genomic analysis. These regions consisted of the 16S mitochondrial rRNA gene (16S) and the cytochrome c oxidase subunit I (COI), which were queried in GenBank (Clark et al., 2016) and for COI also in the Barcode of Life Data Systems (BOLD)[92]. Moreover, for BOLD, species-level identifications were classified into three categories of taxonomic congruence following Leite et al. (2020)[62]: *discordant*, where a single BIN is shared by more than one species; *multiple BIN*, where a single species is assigned to more than one BIN; and *concordant*, where a single species is exclusively associated with one BIN. This classification enables the evaluation of the correspondence between genetic clusters (BINs) and species-level identifications. BIN assignment depends on multiple factors related to the quality, quantity, and characteristics of the available genetic data. Key determinants include sequence quality and length (> 500 base pairs with < 1% ambiguous nucleotides), taxonomic coverage, availability of essential metadata (such as voucher specimens and geospatial information), genetic divergence and accuracy of taxonomic identifications. These factors reflect the robustness of the genetic data and the taxonomic framework for each species, ensuring the effective use of BINs in biodiversity studies.

DNA extraction, amplification, sequencing and barcoding assignment

DNA was extracted from a small tissue sample of each individual using the DNeasy Blood & Tissue DNA Kit (Qiagen Iberia S.L.; Barcelona, Spain) following the manufacturer's instructions with minor modifications. Specimens with calcareous structures were homogenized with a handheld rotor-stator prior to incubation to facilitate lysis. Samples were digested overnight with 250 µg of RNase A (Macherey-Nagel™; Düren, Germany), then centrifuged and transferred into new tubes to remove debris. The concentration and purity of the isolated DNA was assessed using a NanoDrop™ One C spectrophotometer (Thermo Fisher Scientific, Inc.; Waltham, US-MA), and integrity was checked by 1% agarose gel electrophoresis stained with Red Safe™ Nucleic Acid Staining Solution (iNtRON Biotechnology; Seoul, Korea).

For species identification, primers targeting the 16S and COI mitochondrial regions, which have been extensively validated in barcoding studies, were employed[93,94]. These primers provide complementary phylogenetic resolution, enable standardized comparisons and are broadly applicable across marine invertebrate taxa. The targeted regions also allow the development of shorter amplicons suitable for eDNA analyses, facilitating scalable benthic assessments. DNA amplifications were performed using 16Sar-L (5'-CGCCTGTTTATCAAAAACAT-3') and 16Sbr-H (5'-CCGGTCTGAACTCAGATCACGT-3') for 16S rRNA and jgLCO1490 (5'-TITCIACIAAYCAYAARGAYATTGG-3') and jgHCO2198 (5'-TAIACYTCIGGRTGICCRAARAAYCA-3') for COI. PCR reactions consisted of 40 µL containing 1x Buffer A (Nippon Genetics; Düren, Germany), 0.45 mM dNTPs (Thermo Fisher Scientific, Inc.; Waltham, US-MA), 2.5 mM $MgCl_2$ (Nippon Genetics; Düren, Germany), 0.4 µM of each primer, 200 ng BSA (Thermo Fisher Scientific, Inc.; Waltham, US-MA), 75 ng of DNA template and 0.5 U of FastGene® Taq DNA Polymerase (Nippon Genetics; Düren, Germany). Reactions included positive and negative controls, as well as three technical replicates per sample, and were run in a Bio-Rad CFX Connect Real-Time PCR thermal cycler (Bio-Rad Laboratories, Inc.; Hercules, US-CA). The amplification conditions for 16S were as follows: an initial denaturation at 95 ºC for 5 min, followed by 35 cycles of denaturation at 94 ºC for 1 min, annealing at 55 ºC for 1 min and extension at 72 ºC for 2 min, with a final extension at 72 ºC for 10 min. The thermocycling profile for COI amplification consisted of an initial denaturation at 95 ºC for 5 min, followed by 40 cycles of denaturation at 95 ºC for 1 min, annealing at 48 ºC for 1 min and extension at 72 ºC for 1 min, with a final extension at 72 ºC for 5 min. PCR products were examined on 2% agarose gels stained with Red Safe™ Nucleic Acid Staining Solution (iNtRON Biotechnology; Seoul, Korea) and visualized using Gel Doc™ XR + (Bio-Rad, Hercules, CA). For PCR products showing multiple bands, amplicons of the target size (~ 710 pb for COI and ~ 620 pb for 16S rDNA) were excised and purified using a QIAquick Gel Extraction Kit (Qiagen Iberia S.L.; Barcelona, Spain). Purified PCR products were bidirectionally sequenced on an ABI3730XL DNA analyzer (Applied Biosystems Inc., Foster City, CA) at Macrogen Spain (Madrid).

The forward and reverse sequences were reviewed, manually assembled, trimmed and edited using Geneious Prime® 2023.2.1 (https://www.geneious.com). The resulting consensus sequences for COI were amino-acid translated to check for premature stop codons, which could indicate introns or pseudogenes. For barcoding assignment, consensus sequences were searched against GenBank, and the BOLD database using BLASTn[95] and the BOLD identification engine[92], respectively. A sequence identity threshold of 97% and an E- value of 0 were applied[96], and the best BLAST hit or top BOLD match was used for taxonomic assignment.

Integrated morphology and DNA barcoding taxonomy

Taxonomic assignments obtained from DNA barcoding were systematically compared with morphological identifications to evaluate the concordance between both approaches. For specimens showing discordant results or low taxonomic resolution, a stepwise review was conducted. First, morphological identifications were re-examined, and DNA sequences were re-checked for quality and potential errors.

Only samples with successful technical amplification for at least one marker (COI or 16S) were retained for analysis. Records were considered “congruent” if the morphological identification matched at least one molecular identification (BLAST for COI or 16S, or BOLD for COI) at ≥ 97% sequence similarity, this threshold was selected as a conservative compromise between species-level resolution and database limitations. In cases of multiple top matches, priority was given to records with voucher specimens and complete metadata. “Non-congruent” records were further classified based on the availability of reference sequences in public databases. If reference sequences existed for at least one marker, the record was labelled “Pending verification”, potentially indicating misidentification or open taxonomy (e.g., cf.). If no reference sequences were available (i.e., missing entries or Low Taxonomic Resolution, LTR), the record

was flagged as a “Potential new species”. This framework allowed simultaneous identification of congruent records, discrepancies, and candidate taxa for further taxonomic investigation.

Persistent discrepancies were resolved by expert taxonomists, who reassessed specimens using both molecular and morphological evidence. Individuals were examined under a stereo zoom microscope Leica S APO Stereo Zoom and identified to the lowest possible taxonomic level based on morphological characters, using appropriate identification keys. To ensure consistency in taxonomic nomenclature across morphological and molecular assignments, taxonomic nomenclature was standardized following the World Register of Marine Species (WoRMS) (WoRMS, 2026), using the WoRMS online taxon match tool (https://www.marinespecies.org/).

Data analysis

The data analysis was conducted using RStudio 2024.12.1 + 563 "Kousa Dogwood" Release for Ubuntu Focal, running R version 4.4.3 (2025-02-28, "Trophy Case"). Analyses were performed on a Linux x86_64 platform, and visualizations were carried out using the *tidyverse* package for data manipulation and plotting[97], along with *viridis* for colour-blind-friendly palettes[98].

The analysis involved a comprehensive evaluation of DNA barcoding effectiveness for marine specimen identification using multiple analytical approaches. First, taxonomic labels were standardized and parsed using regular expressions to enable automated counting and summarisation. Sequence availability across reference databases (GenBank COI, GenBank 16S, and BOLD COI) was then examined, and the proportion of specimens with available molecular reference data was calculated. The core barcoding analysis evaluated molecular identification success using similarity thresholds to determine assignment rates against reference databases. A bidimensional success matrix was subsequently generated to distinguish between technical success (PCR amplification and sequencing) and database assignment success (matching against reference sequences). Finally, taxonomic stratification by phylum and order was performed to identify biases in barcoding performance, followed by a comparative analysis across molecular markers (COI vs. 16S) and reference databases (GenBank vs. BOLD) to assess the relative effectiveness of each approach for marine biodiversity assessment.

Declarations

Acknowledgements

The authors would like to thank the collaboration and work of the Flemish Cap 2022 research survey leader team from the Spanish Institute of Oceanography (IEO-CSIC), the rest of the scientific staff from the Institute of Marine Research (IIM-CSIC) and the Instituto Português do Mar e da Atmosfera (IPMA), as well as the Vizconde de Eza’s crew, for their performance during the scientific survey. MP would like to thank Dr. María López-Acosta for generously sharing her network of contacts, which enabled us to establish a collaboration that substantially enhanced the quality of this work. A preliminary analysis of these data was presented at the ICES Annual Congress 2023, both MP and NVA would like to thank the feedback received during the conference.

Author Contribution

MP: Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing; **BA:** Methodology, Formal analysis, Data curation, Writing – review & editing; **NVA:** Investigation, Methodology, Writing – review & editing; **DCA:** Resources**,** Methodology; **FGC:** Resources, Methodology, Writing – review & editing; **CD:** Methodology, Writing – review & editing; **AS:** Methodology, Writing – review & editing; **CLH:** Methodology; **MC:** Methodology, Writing – review & editing; **FSR:** Conceptualization, Investigation, Funding acquisition, Project administration, Resources, Writing – review & editing; **LC:** Conceptualization, Investigation, Supervision, Validation, Writing – original draft, Writing – review & editing

Data Availability

All sequences have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB81803 (https://www.ebi.ac.uk/ena/browser/view/PRJEB81803) and in the Barcode of Life Data Systems (BOLD) under dataset DS-VMEFC (dx.doi.org/10.5883/DS-VMEFC).

Voucher specimens for all collected taxa are deposited at the Institute of Marine Research (IIM-CSIC) in Vigo, Spain, and are available for further research.

Declaration of Interest statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This work was carried out within the framework of the research project “Fisheries sustainability and protection of the biodiversity of vulnerable marine ecosystems, established under the agreement signed in September 2021 between the Spanish Ministry of Agriculture, Fisheries and Food and the Spanish National Research Council, to promote fisheries research as a basis for sustainable fisheries management, and funded by the European Union – NextGeneration EU. The EU Bottom Trawl Survey on Flemish Cap (NAFO Division 3M) was co-financed by the EU through the European Maritime, Fisheries

and Aquaculture Fund (EMFAF) within the Spanish National Programme for the collection, management and use of data in the fisheries sector, and to support scientific advice related to the Common Fisheries Policy. **MP** was supported by Grant FJC2021-047881-I funded by MCIN/AEI/10.13039/501100011033 and by the European Union NextGenerationEU/PRTR. **NVA** was supported by Grant PTA2021-020507-I funded by MCIN/AEI/10.13039/501100011033 and by the ESF+.

References

1. Paulus, E. Shedding Light on Deep-Sea Biodiversity—A Highly Vulnerable Habitat in the Face of Anthropogenic Change. *Front Mar. Sci* **8**, (2021).
2. Ramirez-Llodra, E. et al. Deep, diverse and definitely different: unique attributes of the world's largest ecosystem. *Biogeosciences* **7**, 2851–2899 (2010).
3. Levin, L. A. Le Bris, N. The deep ocean under climate change. *Science* **350**, 766–768 (2015).
4. Thurber, A. R. et al. Ecosystem function and services provided by the deep sea. *Biogeosciences* **11**, 3941–3963 (2014).
5. FAO. International Guidelines for the Management of Deep-sea Fisheries in the High Seas/Directives internationales sur la gestion de la pêche profonde en haute mer/ Directrices Internacionales para la Ordenación de las Pesquerías de Aguas Profundas en Alta Mar. (FAO, Rome, Italy, (2009).
6. Roberts, J. M., Cold-Water & Corals *The Biology and Geology of Deep-Sea Coral Habitats* (Cambridge University Press, 2009).
7. Roberts, J. M. & Cairns, S. D. Cold-water corals in a changing ocean. *Curr. Opin. Environ. Sustain.* **7**, 118–126 (2014).
8. Murillo, F. J., Serrano, A., Kenchington, E. & Mora, J. Epibenthic assemblages of the Tail of the Grand Bank and Flemish Cap (northwest Atlantic) in relation to environmental parameters and trawling intensity. *Deep Sea Res. Part I*. **109**, 99–122 (2016).
9. Pham, C. K. et al. Removal of deep-sea sponges by bottom trawling in the Flemish Cap area: conservation, ecology and economic assessment. *Sci. Rep.* **9**, 15843 (2019).
10. Åström, E. K. L., Sen, A., Carroll, M. L. & Carroll, J. Cold Seeps in a Warming Arctic: Insights for Benthic Ecology. *Front Mar. Sci* **7**, (2020).
11. Menini, E. & Van Dover, C. L. An atlas of protected hydrothermal vents. *Mar. Policy*. **108**, 103654 (2019).
12. Thompson, A. B. & Reid, K. *Review of the Implementation of the International Guidelines for the Management of Deep-Sea Fisheries in the High Seas* (Food & Agriculture Org. [Author], 2024).
13. Griffiths, J. R. et al. The importance of benthic–pelagic coupling for marine ecosystem functioning in a changing world. *Glob. Change Biol.* **23**, 2179–2196 (2017).
14. Buhl-Mortensen, L. et al. Biological structures as a source of habitat heterogeneity and biodiversity on the deep ocean margins. *Mar. Ecol.* **31**, 21–50 (2010).
15. Santín, A. et al. Mediterranean Coral Provinces as a Sponge Diversity Reservoir: Is There a Mediterranean Cold-Water Coral Sponge Fauna? *Front Mar. Sci* **8**, (2021).
16. Coleman, F. C. & Williams, S. L. Overexploiting marine ecosystem engineers: potential consequences for biodiversity. *Trends Ecol. Evol.* **17**, 40–44 (2002).
17. Sciberras, M. et al. Response of benthic fauna to experimental bottom fishing: A global meta-analysis. *Fish Fish.* **19**, 698–715 (2018).
18. Baco, A. R. et al. Towards a scientific community consensus on designating Vulnerable Marine Ecosystems from imagery. *PeerJ* **11**, e16024 (2023).
19. Wang, S. et al. Quantifying the effects of fragmentation of connectivity networks of deep-sea vulnerable marine ecosystems. *Divers. Distrib.* **30**, e13824 (2024).
20. Kenchington, E. et al. Kernel Density Surface Modelling as a Means to Identify Significant Concentrations of Vulnerable Marine Ecosystem Indicators. *PLOS ONE*. **9**, e109365 (2014).
21. Gros, C. et al. Identifying vulnerable marine ecosystems: an image-based vulnerability index for the Southern Ocean seafloor. *ICES J. Mar. Sci.* **80**, 972–986 (2023).
22. Morato, T. et al. A Multi Criteria Assessment Method for Identifying Vulnerable Marine Ecosystems in the North-East Atlantic. *Front Mar. Sci* **5**, (2018).
23. Ovenden, J. R., Berry, O., Welch, D. J., Buckworth, R. C. & Dichmont, C. M. Ocean's eleven: a critical evaluation of the role of population, evolutionary and molecular genetics in the management of wild fisheries. *Fish Fish.* **16**, 125–159 (2015).
24. Ruppert, K. M., Kline, R. J. & Rahman, M. S. Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. *Global Ecol. Conserv.* **17**, e00547 (2019).
25. Antich, A. et al. Marine biomonitoring with eDNA: Can metabarcoding of water samples cut it as a tool for surveying benthic communities? *Mol. Ecol.* **30**, 3175–3188 (2021).
26. Clark, A. J. et al. Cost-effort analysis of Baited Remote Underwater Video (BRUV) and environmental DNA (eDNA) in monitoring marine ecological communities. *PeerJ* **12**, e17091 (2024).
27. Thomsen, P. F. & Willerslev, E. Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biol. Conserv.* **183**, 4–18 (2015).
28. Kasmi, Y. et al. Atlantic cod (Gadus morhua) assessment approaches in the North and Baltic Sea: A comparison of environmental DNA analysis versus bottom trawl sampling. *Front Mar. Sci* **10**, (2023).
29. Shelton, A. O. et al. Environmental DNA provides quantitative estimates of Pacific hake abundance and distribution in the open ocean. *Proc. Biol. Sci.* **289**, 20212613 (2022).
30. Hestetun, J. T. et al. Significant taxon sampling gaps in DNA databases limit the operational use of marine macrofauna metabarcoding. *Mar. Biodivers.* **50**, 70 (2020).

31. Marques, V. et al. GAPeDNA: Assessing and mapping global species gaps in genetic databases for eDNA metabarcoding. *Divers. Distrib.* **27**, 1880–1892 (2021).
32. Zhou, Y. et al. Evaluation of DNA barcoding reference databases for marine species in the western and central Pacific Ocean. *PeerJ* **13**, e19674 (2025).
33. Chen, E. Y. S. Often Overlooked: Understanding and Meeting the Current Challenges of Marine Invertebrate Conservation. *Front Mar. Sci* **8**, (2021).
34. Taylor, M. L. & Roterman, C. N. Invertebrate population genetics across Earth's largest habitat: The deep-sea floor. *Mol. Ecol.* **26**, 4872–4896 (2017).
35. Benestan, L. Population Genomics Applied to Fishery Management and Conservation. in Population Genomics: Marine Organisms (eds (eds Oleksiak, M. F. & Rajora, O. P.) 399–421 (Springer International Publishing, Cham, doi:10.1007/13836_2019_66. (2020).
36. Valenzuela-Quiñonez, F. How fisheries management can benefit from genomics? *Brief. Funct. Genomics*. **15**, 352–357 (2016).
37. Sacau, M. M., Neves, B. M. & Durán-Muñoz, P. New preliminary data on VME encounters in NAFO Regulatory Area (Div. 3M) from EU-Spain and Portugal Groundfish Surveys (2020) and Canadian surveys. (2020). https://digital.csic.es/handle/10261/327794
38. NAFO, Conservation & Measures, E. 200 (2023). https://www.nafo.int/Portals/0/PDFs/COM/2023/comdoc23-01.pdf (2023).
39. Durán Muñoz, P., Sacau, M., García-Alegre, A. & Román, E. Cold-water corals and deep-sea sponges by-catch mitigation: Dealing with groundfish survey data in the management of the northwest Atlantic Ocean high seas fisheries. *Mar. Policy*. **116**, 103712 (2020).
40. Vázquez, A., Casas-Sánchez, J. M. & Alpoim, R. Protocols of the EU bottom trawl survey of Flemish Cap. (2014). https://doi.org/10.2960/S.v46.m1 doi:10.2960/S.v46.m1.
41. NAFO. Report of the Scientific Council Working Group on Ecosystem Science and Assessment. (2019). 18 ± 29 November 2019, Dartmouth, Nova Scotia, Canada. NAFO SCS Doc. 19/25.
42. NAFO. Report of the Scientific Council Working Group on Ecosystem Science and Assessment. 15–24 (2022). November 2022, Dartmouth, Nova Scotia, Canada. NAFO SCS Doc. 22/25.
43. Keck, F., Couton, M. & Altermatt, F. Navigating the seven challenges of taxonomic reference databases in metabarcoding analyses. *Mol. Ecol. Resour.* **23**, 742–755 (2023).
44. Emerson, B. C. Delimiting Species—Prospects and Challenges for DNA Barcoding. *Mol. Ecol.* **34**, e17677 (2025).
45. Kaiser, S. et al. Editorial: Recent and emerging innovations in deep-sea taxonomy to enhance biodiversity assessment and conservation. *Front Mar. Sci* **9**, (2022).
46. Meier, R. et al. Dark taxonomy: A new protocol for overcoming the taxonomic impediments for dark taxa and broadening the taxon base for biodiversity assessment. *Cladistics* **41**, 223–238 (2025).
47. Dell'Anno, A., Carugati, L., Corinaldesi, C., Riccioni, G. & Danovaro, R. Unveiling the Biodiversity of Deep-Sea Nematodes through Metabarcoding: Are We Ready to Bypass the Classical Taxonomy? *PLOS ONE*. **10**, e0144928 (2015).
48. Liu, J. & Zhang, H. Combining Multiple Markers in Environmental DNA Metabarcoding to Assess Deep-Sea Benthic Biodiversity. *Front Mar. Sci* **8**, (2021).
49. McClenaghan, B. et al. Harnessing the power of eDNA metabarcoding for the detection of deep-sea fishes. *PLOS ONE*. **15**, e0236540 (2020).
50. Beazley, L. I., Anstey, L. J. & Kenchington, E. R. Summary of the Location of VME Indicators on the Flemish Cap Slope Based on in Situ Benthic Imagery Analysis. 1–18 (2013). https://www.researchgate.net/publication/299538352_Summary_of_the_location_of_VME_indicators_on_the_Flemish_Cap_slope_based_on_in_situ_bent
51. Downie, A. L. et al. Reconstructing baselines: use of habitat suitability modelling to predict pre-fishing condition of a Vulnerable Marine Ecosystem. *ICES J. Mar. Sci.* **78**, 2784–2796 (2021).
52. Murillo, F. J., Weigel, B., Bouchard Marmen, M. & Kenchington, E. Marine epibenthic functional diversity on Flemish Cap (north-west Atlantic)—Identifying trait responses to the environment and mapping ecosystem functions. *Divers. Distrib.* **26**, 460–478 (2020).
53. Murillo, F. J. et al. Mapping benthic ecological diversity and interactions with bottom-contact fishing on the Flemish Cap (northwest Atlantic). *Ecol. Ind.* **112**, 106135 (2020).
54. McClain, C. R. & Schlacher, T. A. On some hypotheses of diversity of animal life at great depths on the sea floor. *Mar. Ecol.* **36**, 849–872 (2015).
55. Ramirez-Llodra, E. et al. The emerging picture of a diverse deep Arctic Ocean seafloor: From habitats to ecosystems. *Elementa: Sci. Anthropocene*. **12**, 00140 (2024).
56. Ravara, A., Ramos, D., Teixeira, M. A. L., Costa, F. O. & Cunha, M. R. Taxonomy, distribution and ecology of the order Phyllodocida (Annelida, Polychaeta) in deep-sea habitats around the Iberian margin. *Deep Sea Res. Part II*. **137**, 207–231 (2017).
57. Liu, J. & Zhang, H. DNA barcoding for species identification in deep-sea clams (Mollusca: Bivalvia: Vesicomyidae). *Mitochondrial DNA Part. A*. **29**, 1165–1173 (2018).
58. Taite, M., Dillon, L., Strugnell, J. M., Drewery, J. & Allcock, A. L. DNA Barcoding Reveals Unexpected Diversity of Deep-Sea Octopuses in the North-East Atlantic. *Biology and Environment: Proceedings of the Royal Irish Academy* 123, 1–12 (2023).
59. Raupach, M. J. et al. The Application of DNA Barcodes for the Identification of Marine Crustaceans from the North Sea and Adjacent Regions. *PLOS ONE*. **10**, e0139421 (2015).
60. Wörheide, G., Erpenbeck, D. & Menke, C. The Sponge Barcoding Project: Aiding in the identification and description of poriferan taxa. *Porifera Research: Biodivers. Innov. Sustain.* 123–128 (2007).
61. Weigand, H. et al. DNA barcode reference libraries for the monitoring of aquatic biota in Europe: Gap-analysis and recommendations for future work. *Sci. Total Environ.* **678**, 499–524 (2019).

62. Leite, B. R. et al. Gap-analysis and annotated reference library for supporting macroinvertebrate metabarcoding in Atlantic Iberia. *Reg. Stud. Mar. Sci.* **36**, 101307 (2020).
63. Wangensteen, O. S., Palacín, C., Guardiola, M. & Turon, X. DNA metabarcoding of littoral hard-bottom communities: high diversity and database gaps revealed by two molecular markers. *PeerJ* **6**, e4705 (2018).
64. Bell, J. J. Contrasting patterns of species and functional composition of coral reef sponge assemblages. *Mar. Ecol. Prog. Ser.* **339**, 73–81 (2007).
65. Grupstra, C. G. B. et al. Integrating cryptic diversity into coral evolution, symbiosis and conservation. *Nat. Ecol. Evol.* **8**, 622–636 (2024).
66. Todd, P. A. Morphological plasticity in scleractinian corals. *Biol. Rev. Camb. Philos. Soc.* **83**, 315–337 (2008).
67. Bridges, A. E. H. & Howell, K. L. Prioritisation of ocean biodiversity data collection to deliver a sustainable ocean. *Commun. Earth Environ.* **6**, 473 (2025).
68. Claver, C., Canals, O., de Amézaga, L. G. & Mendibil, I. Rodriguez-Ezpeleta, N. An automated workflow to assess completeness and curate GenBank for environmental DNA metabarcoding: The marine fish assemblage as case study. *Environ. DNA*. **5**, 634–647 (2023).
69. Teixeira, M. A. L., Aylagas, E., Pearman, J. K. & Carvalho, S. Gaps and Data Ambiguities in DNA Reference Libraries: A Limiting Factor for Molecular-Based Biodiversity Assessments Using Annelids as a Case Study. *Ecol. Evol.* **15**, e71544 (2025).
70. Barua, A., Afrin, T., Akhand, A. A. & Ahmed, M. S. Molecular characterization and phylogenetic analysis of crabs (Crustacea: Decapoda: Brachyura) based on mitochondrial COI and 16S rRNA genes. *Conserv. Genet. Resour.* **13**, 291–301 (2021).
71. Ali, A. J., Mohammed Kaleem Arshan, H. & Praba, M. L. Kaleemullah Khan, B. Biological identification of *Eudistoma* species (Polycitoridae: Ascidiacea) from Gulf of Mannar, India through COI gene sequences. *Gene Rep.* **24**, 101222 (2021).
72. Kongsrud, J. A., Eilertsen, M. H., Alvestad, T., Kongshavn, K. & Rapp, H. T. New species of Ampharetidae (Annelida: Polychaeta) from the Arctic Loki Castle vent field. *Deep Sea Res. Part II*. **137**, 232–245 (2017).
73. Vargas, S. et al. Barcoding Sponges: An Overview Based on Comprehensive Sampling. *PLoS ONE*. **7**, e39345 (2012).
74. Hoareau, T. B. & Boissin, E. Design of phylum-specific hybrid primers for DNA barcoding: addressing the need for efficient COI amplification in the Echinodermata. *Mol. Ecol. Resour.* **10**, 960–967 (2010).
75. Lobo, J. et al. Enhanced primers for amplification of DNA barcodes from a broad range of marine metazoans. *BMC Ecol.* **13**, 34 (2013).
76. Meyer, A. et al. Morphological vs. DNA metabarcoding approaches for the evaluation of stream ecological status with benthic invertebrates: Testing different combinations of markers and strategies of data filtering. *Mol. Ecol.* **30**, 3203–3220 (2021).
77. Shearer, T. L., van Oppen, M. J. H., Romano, S. L. & Wörheide, G. Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). *Mol. Ecol.* **11**, 2475–2487 (2002).
78. Shearer, T. L., Coffroth, M. A. & DNA, B. A. R. C. O. D. I. N. G. Barcoding corals: limited by interspecific divergence, not intraspecific variation. *Mol. Ecol. Resour.* **8**, 247–255 (2008).
79. Wei, Z. F. et al. Molecular phylogenetic relationships based on mitochondrial genomes of novel deep-sea corals (Octocorallia: Alcyonacea): Insights into slow evolution and adaptation to extreme deep-sea environments. *Zool. Res.* **45**, 215–225 (2024).
80. Cárdenas, P. et al. Taxonomy, biogeography and DNA barcodes of Geodia species (Porifera, Demospongiae, Tetractinellida) in the Atlantic boreo-arctic region. *Zool. J. Linn. Soc.* **169**, 251–311 (2013).
81. Nydam, M. L. et al. Phylogenomic and morphological relationships among the botryllid ascidians (Subphylum Tunicata, Class Ascidiacea, Family Styelidae). *Sci. Rep.* **11**, 8351 (2021).
82. Zhang, Q. L. et al. A Phylogenomic Framework and Divergence History of Cephalochordata Amphioxus. *Front Physiol* **9**, (2018).
83. Soest, R. W. M. V. et al. Global Diversity of Sponges (Porifera). *PLOS ONE*. **7**, e35105 (2012).
84. Sigwart, J. D. & Lindberg, D. R. Consensus and Confusion in Molluscan Trees: Evaluating Morphological and Molecular Phylogenies. *Syst. Biol.* **64**, 384–395 (2015).
85. Zrzavý, J., Říha, P., Piálek, L. & Janouškovec, J. Phylogeny of Annelida (Lophotrochozoa): total-evidence analysis of morphology and six genes. *BMC Evol. Biol.* **9**, 189 (2009).
86. Radulovici, A. E. et al. Revision and annotation of DNA barcode records for marine invertebrates: report of the 8th iBOL conference hackathon. *Metabarcoding and Metagenomics* 5, e67862 (2021).
87. Chen, C., Pradillon, F., Diaz-Recio Lorenzo, C. & Alfaro-Lucas, J. M. Integrative taxonomy of two new peltospirid gastropods from Mid-Atlantic Ridge hot vents, including a potentially symbiotic species. *Zool. J. Linn. Soc.* **204**, zlaf055 (2025).
88. Olmo, I. et al. Integrative taxonomy in Syllis prolifera (Annelida, Syllidae): from a unique cosmopolitan species to a complex of pseudocryptic species. *Invert. Systematics* 38, NULL-NULL (2024).
89. Linse, K. Antarctic Marine Biodiversity – a taxonomic crisis? *Antarct. Sci.* **20**, 209–209 (2008).
90. Packer, L., Grixti, J. C., Roughley, R. E. & Hanner, R. The status of taxonomy in Canada and the impact of DNA barcodingThe present review is one of a series of occasional review articles that have been invited by the Editors and will feature the broad range of disciplines and expertise represented in our Editorial Advisory Board. *Can. J. Zool.* **87**, 1097–1110 (2009).
91. Coleman, C. O. Taxonomy in Times of the Taxonomic Impediment – Examples from the Community of Experts on Amphipod Crustaceans. *J. Crustacean Biol.* **35**, 729–740 (2015).
92. Ratnasingham, S., Hebert, P. D. N. & bold The Barcode of Life Data System. *Mol. Ecol. Notes*. **7**, 355–364 (2007). http://www.barcodinglife.org
93. Geller, J., Meyer, C., Parker, M. & Hawk, H. 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**, 851–861 (2013).

94. Palumbi, S. R. Nucleic acids II: the polymerase chain reaction. *Mol. Syst.* 205–247 (1996).
95. Madden, T. *The BLAST Sequence Analysis Tool. The BLAST Sequence Analysis Tool* (National Center for Biotechnology Information (US), 2013).
96. Pappalardo, P. et al. The role of taxonomic expertise in interpretation of metabarcoding studies. *ICES J. Mar. Sci.* **78**, 3397–3410 (2021).
97. Wickham, H. et al. Welcome to the Tidyverse. *J. Open. Source Softw.* **4**, 1686 (2019).
98. Garnier, S. viridis: Colorblind-Friendly Color Maps for R. 0.6.5 (2015). https://doi.org/10.32614/CRAN.package.viridis

Figures

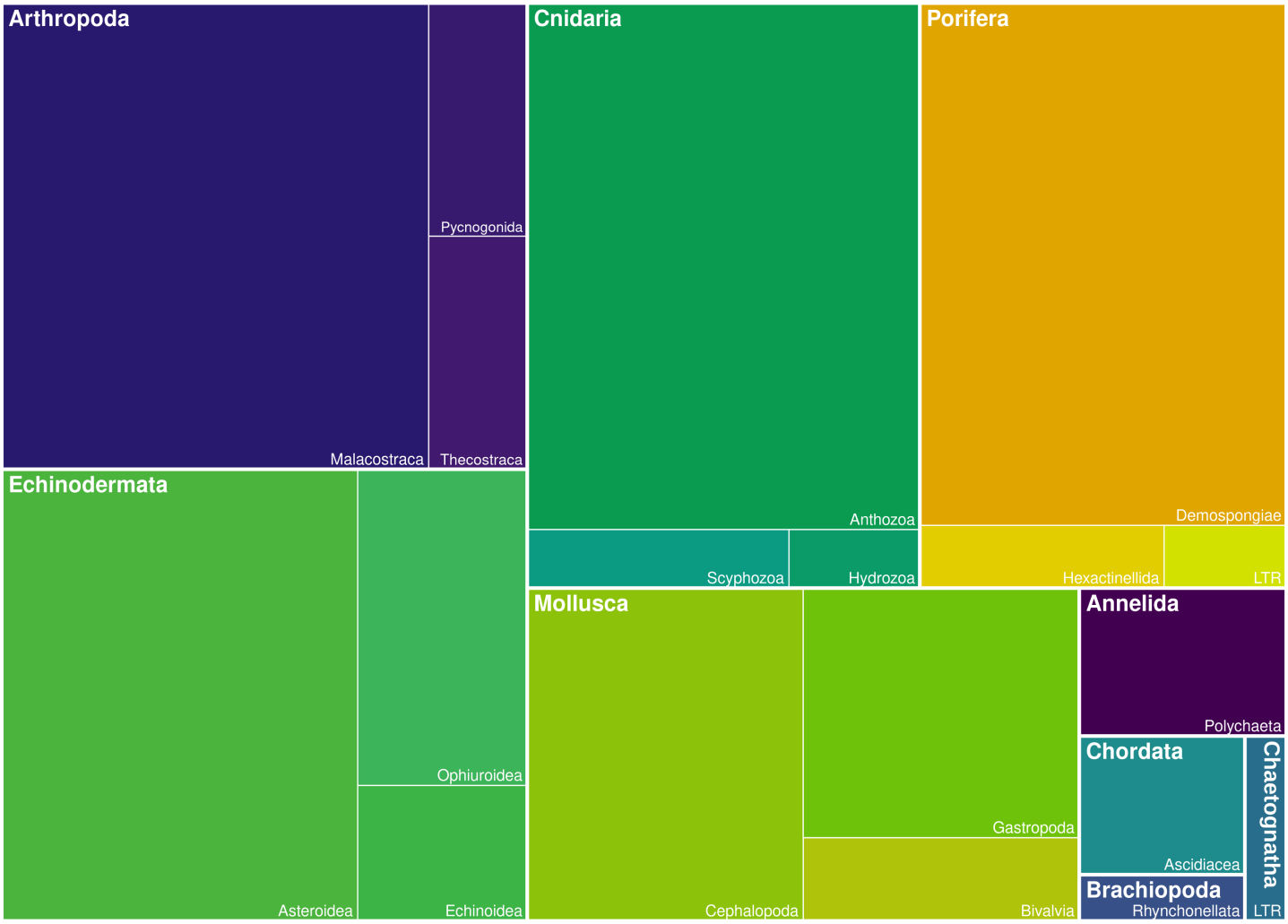


Figure 1

Treemap showing the taxonomic composition of sampled organisms based on the total number of unique taxonomic categories identified morphologically, grouped by phylum. The relative proportion of each class within each phylum is shown. Individuals identified only to phylum (e.g., Porifera and Chaetognatha are recorded as LTR (Low Taxonomic Resolution).

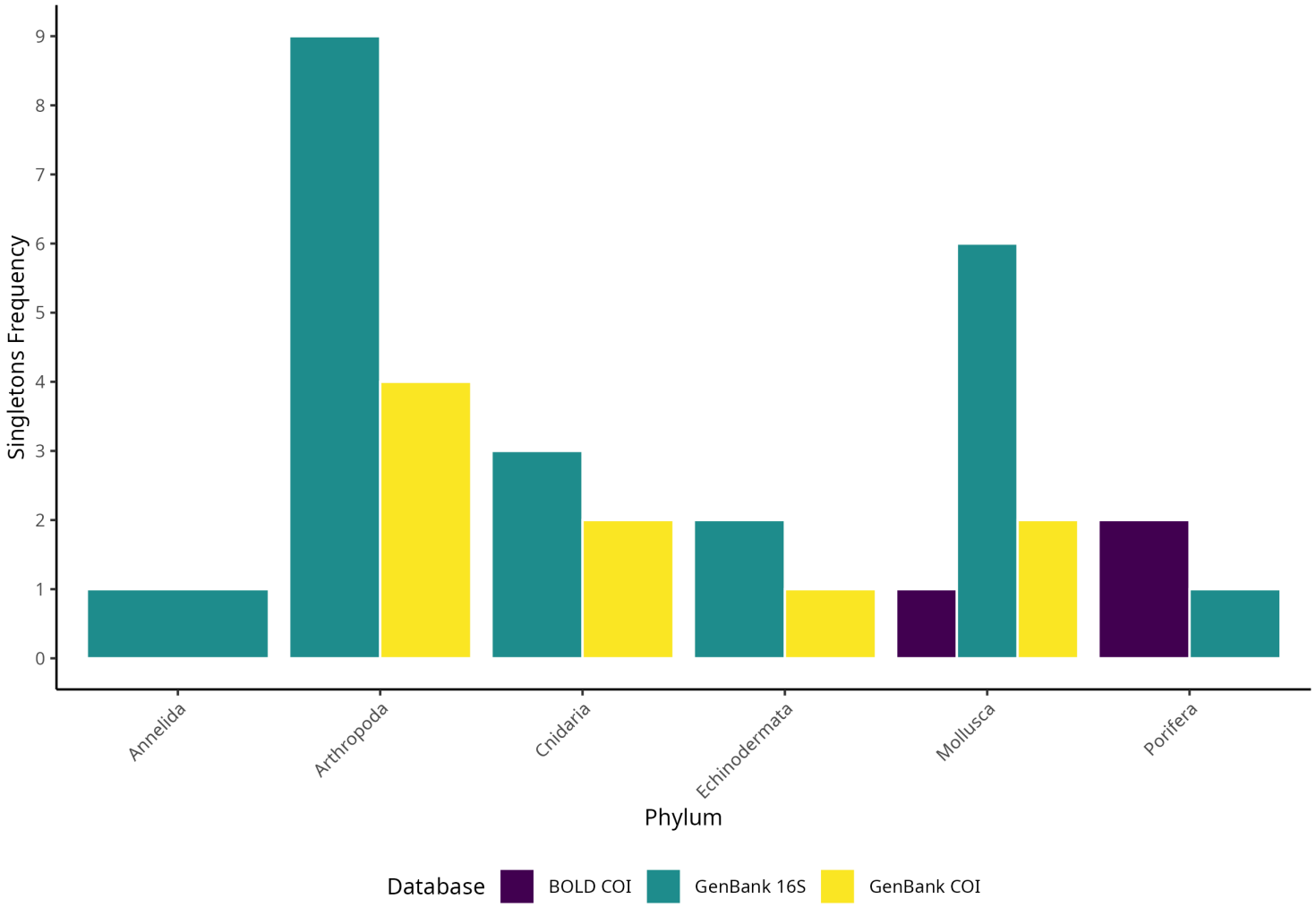


Figure 2

Frequency of singleton species among the 91 morphospecies with reference sequences in public databases.

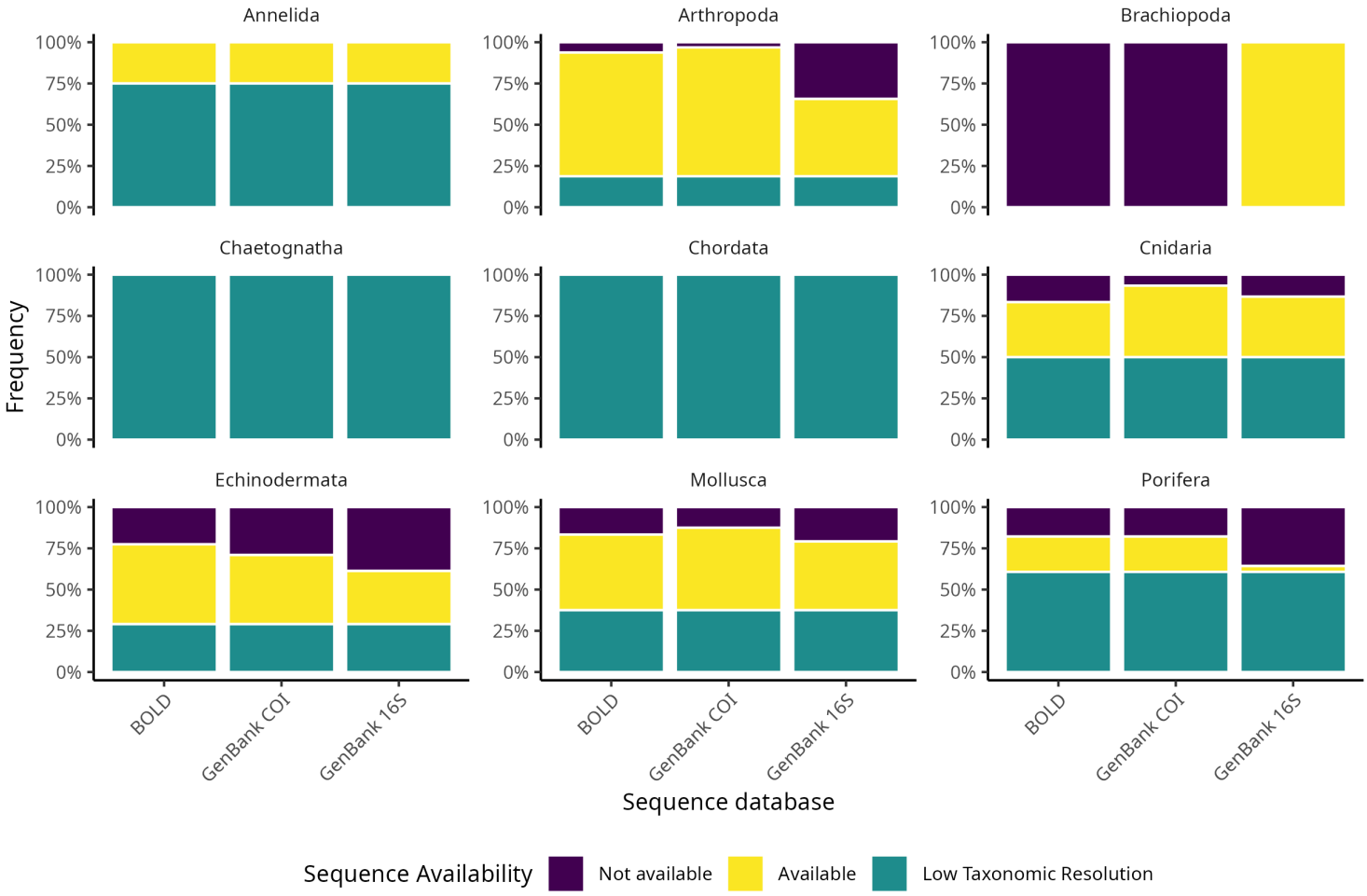


Figure 3

Proportion of morphospecies with representative sequences in public databases (BOLD and GenBank) for COI and 16S, grouped by phylum. "Low Taxonomic Resolution" (LTR) refers to species that could only be assigned to higher taxonomic categories during morphological classification, preventing their inclusion in database searches.

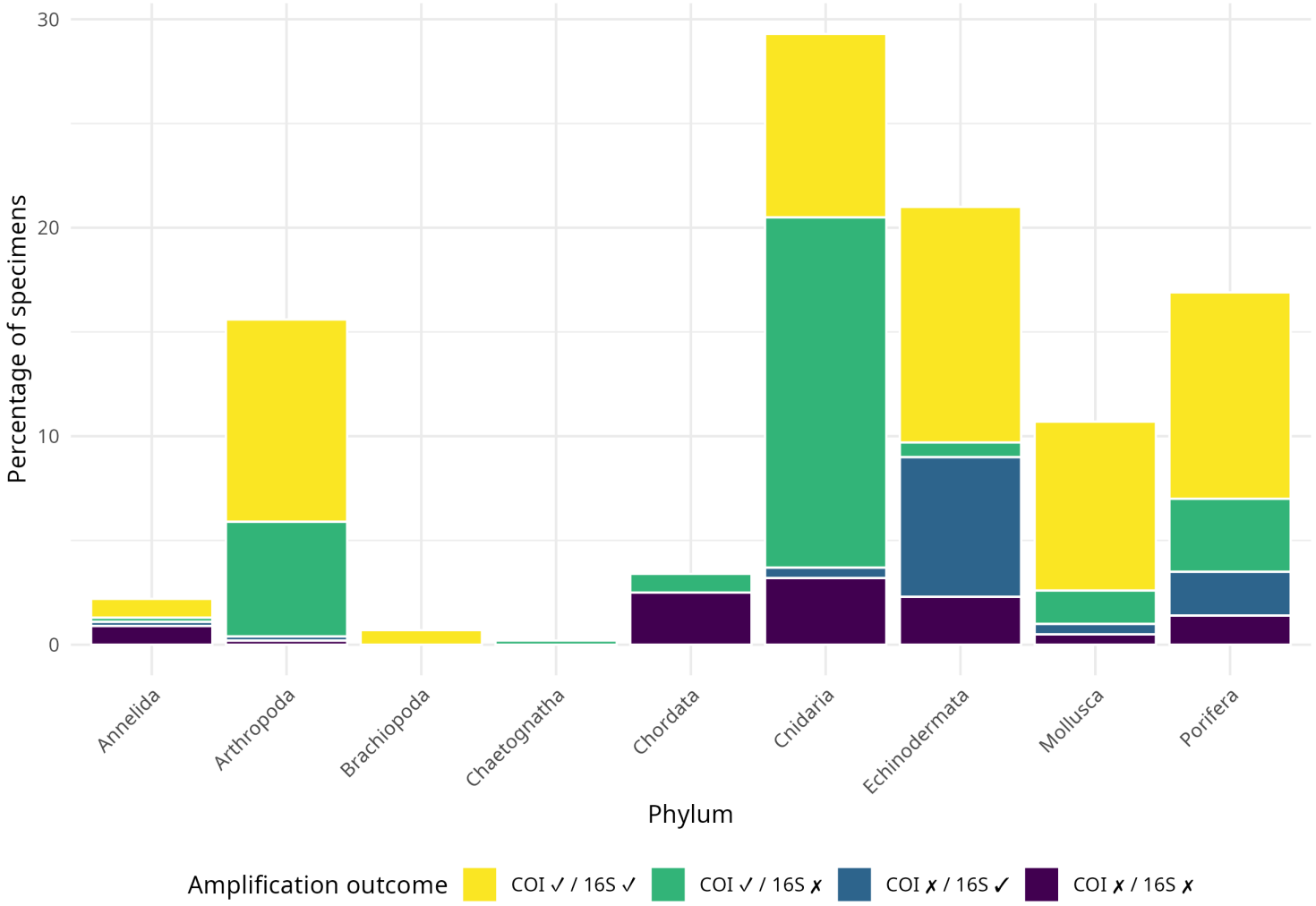


Figure 4

Amplification and sequencing success across taxonomic groups. Bars represent the percentage of specimens within each phylum that were successfully amplified and sequenced for COI and 16S markers.

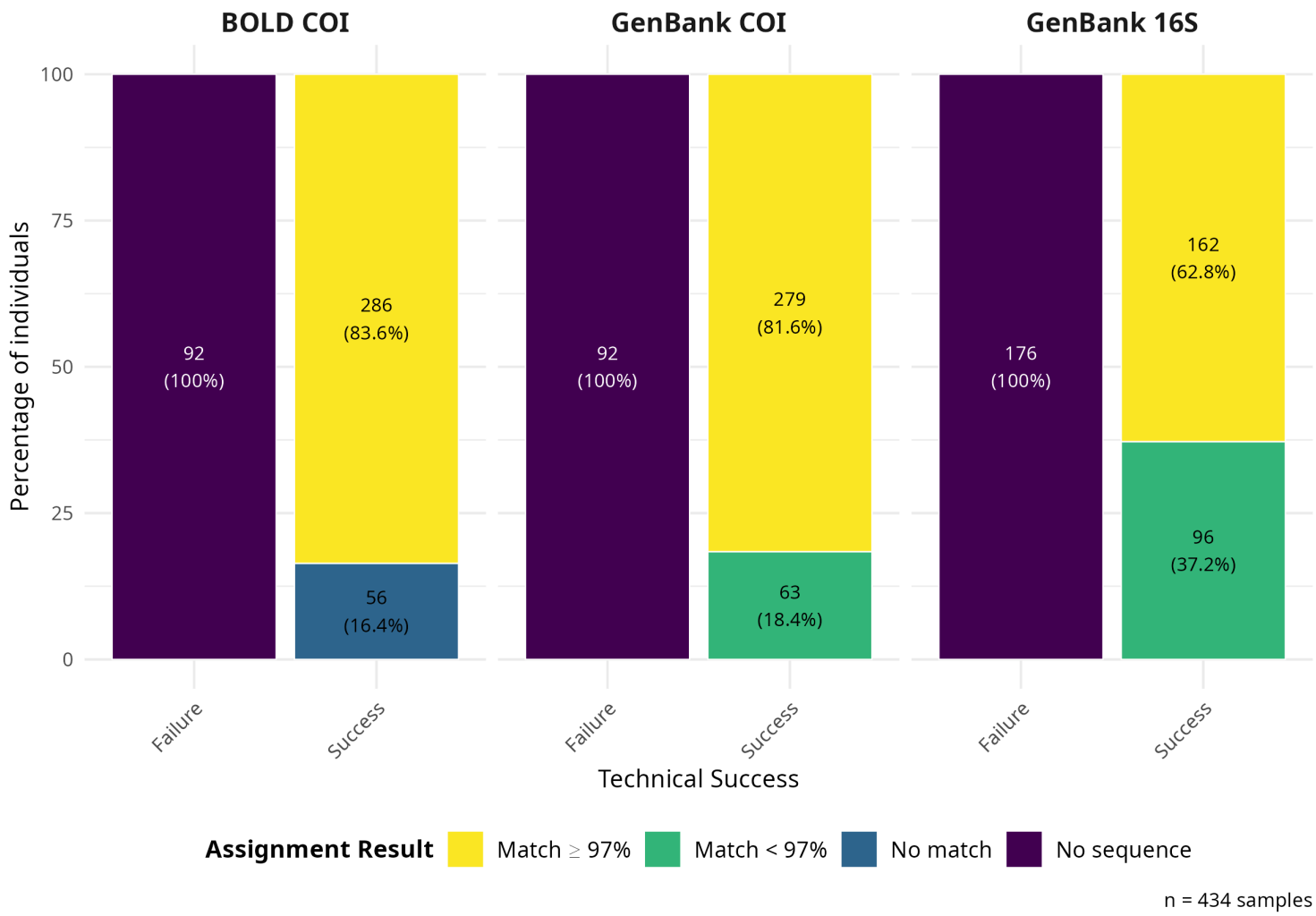


Figure 5

Technical success of samples versus molecular assignment results in BOLD (COI), GenBank (COI), and GenBank (16S). Colours and percentages indicate match levels, where Match ≥ 97: Greater than or equal to 97% pairwise sequence identity; Match < 97: Less than 97% pairwise sequence identity; No match: Samples for which BOLD returns "No match"; No sequence: NA (sequence not available)

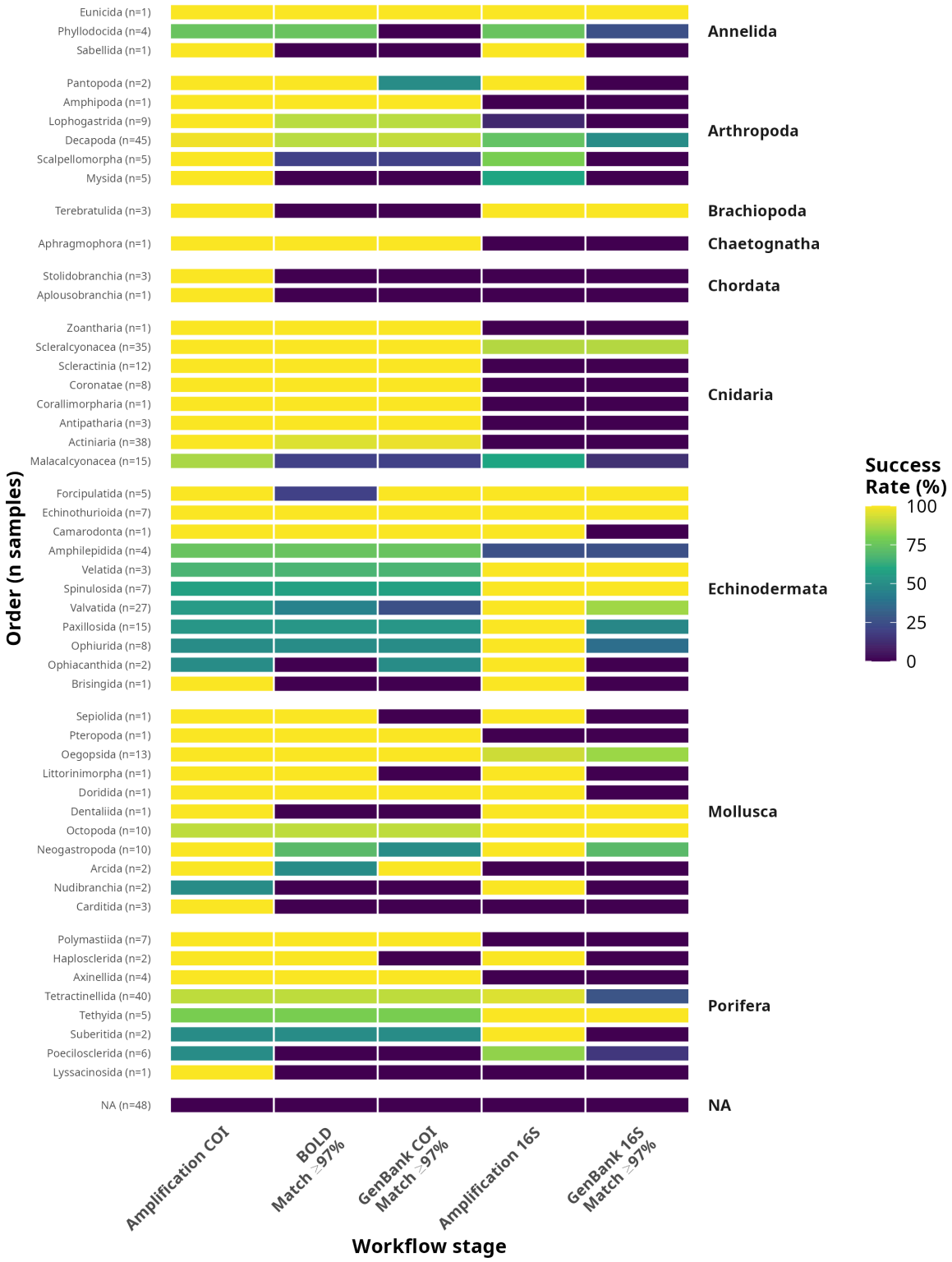


Figure 6

Heatmap of DNA barcoding success (%) across taxonomic orders, grouped by phylum. Colour gradients indicate COI and 16S amplification success, as well as species-level database assignment (≥97% similarity) in BOLD and GenBank. Sample sizes (n) are provided for each order, illustrating order-specific limitations in molecular identification.

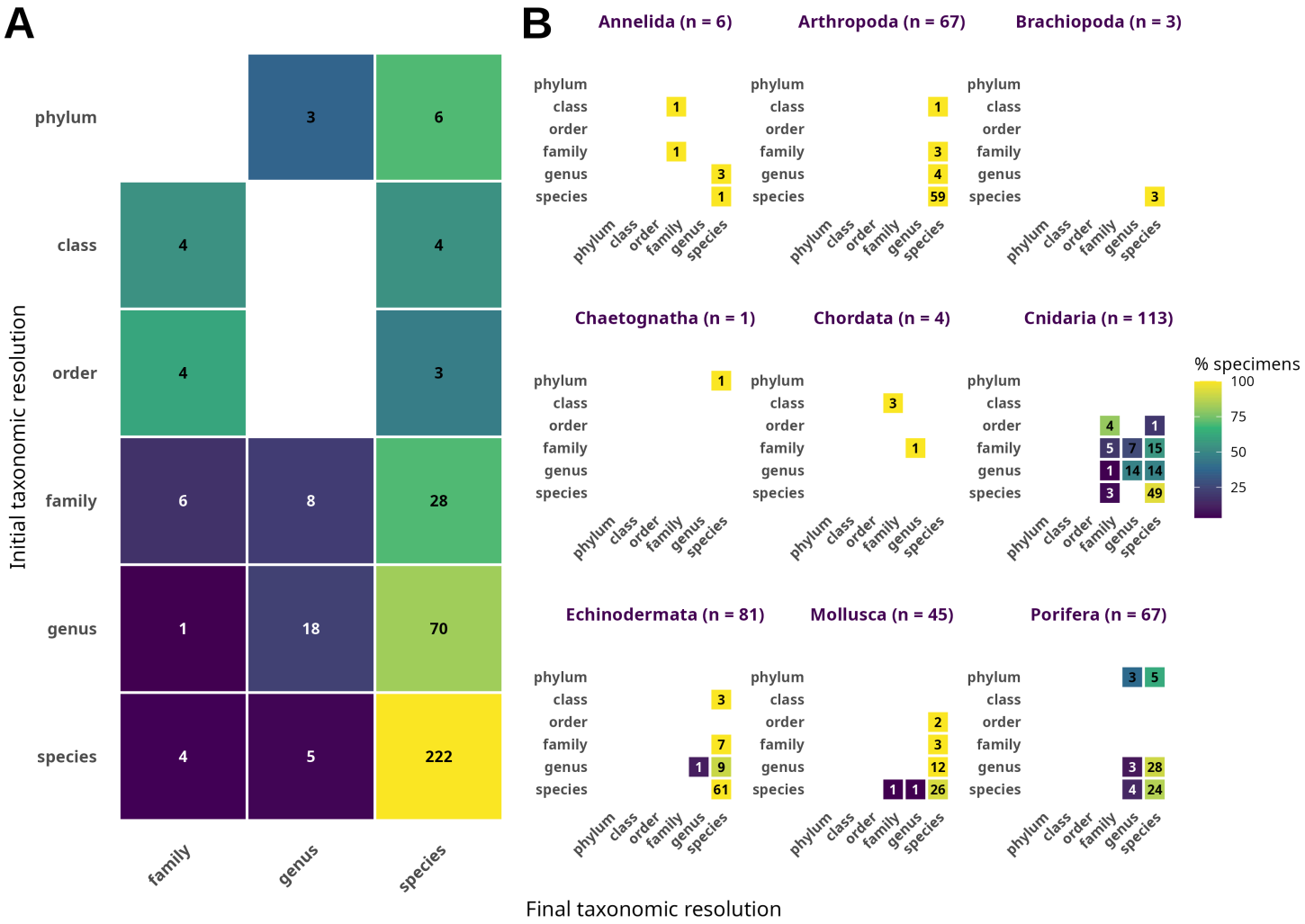


Figure 7

Taxonomic resolution improvement by integrative identification. A) Overall matrix for all specimens with molecular information (n=386). B) Matrices by phylum. Rows represent initial taxonomic resolution and columns represent final resolution. Color intensity indicates the percentage of specimens.

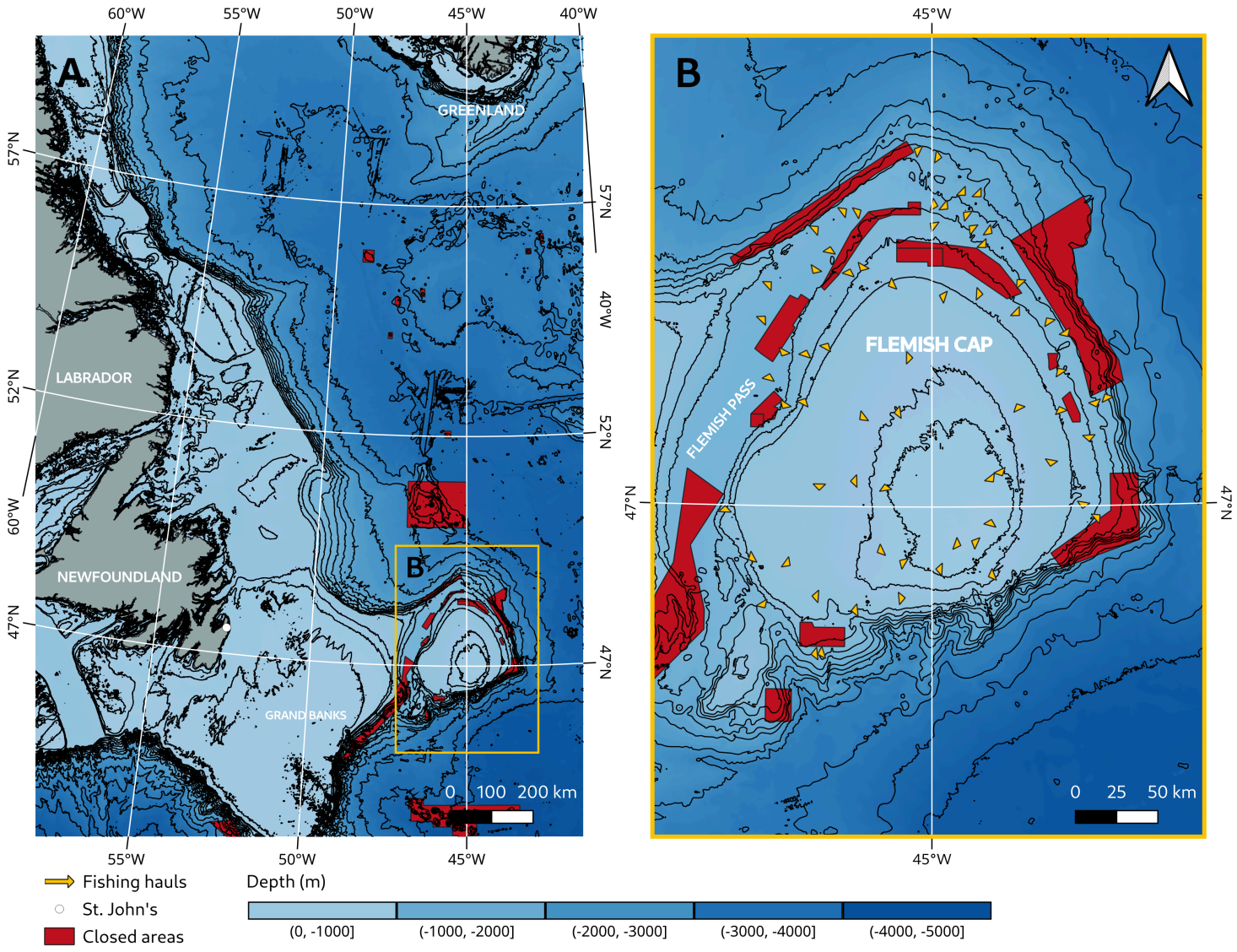


Figure 8

Map of the Flemish Cap area in the Northwest Atlantic: A) Overview of the study area; B) Detailed view showing sampling points (yellow arrows). Areas closed to bottom fishing are outlined in red. Depth is indicated by shading, with darker blue representing greater depth. Map created in QGIS 3.34 (https://www.qgis.org/), using a Lambert Azimuthal Equal Area projection (centered at 54° N, 45° W; datum WGS 84; units in meters).