UNIVERSITY OF GOTHENBURG

DEPARTMENT OF BIOLOGICAL AND ENVIRONMENTAL SCIENCES

MAPPING POLYCHAETE DIVERSITY

Combining novel mega-barcoding and traditional morphology to resolve species diversity and distribution in the genus *Terebellides* (Annelida, Trichobranchidae) in Swedish waters

Amblard-Rambert Paul

Degree project for Master of Science (120 hec) with a major in Biology

BIO717, Degree Project in Evolutionary and behavioural ecology (60 hec)

Second cycle

Semester/year: Spring 2025 - Autumn 2025

Supervisor: Maël Grosse and Arne Nygren (Department of Marine Sciences)

Examiner: Pierre Raoul de Wit (Department of Biological and Environmental Sciences)

Front cover image : Stereomicroscope image of a Terebellides *sp. specimen. Courtesy of Maël Grosse / Department of Marine Sciences (Gothenburg University)*

Table of contents

Abstract

Recent genetic analyses have uncovered a hidden species complex within the genus *Terebellides* Sars, 1835 (Annelida, Trichobranchidae), revealing 27 North East Atlantic species where only seven were previously recognized. In the Skagerrak and Kattegat, ten species are present, eight of them are described. Because *Terebellides* is a dominant taxon along the Swedish West Coast, precise identification is essential for environmental monitoring, yet currently unattainable through morphology alone. This study utilized a novel molecular mega-barcoding approach to assess *Terebellides* diversity and distribution using 2022 and 2024 specimens from the Swedish national environmental monitoring program (PMK). Genomic DNA was extracted from 529 specimens, and a COI mini-barcode was amplified for dual-platform sequencing using Sanger and Oxford Nanopore Technologies (ONT) MinION. Taxonomic assignment was performed via phylogenetic reconstruction against reference clades established through prior multi-marker species delimitation. Morphological verification using light microscopy was conducted on undescribed lineages to identify diagnostic characters, including branchial shape and uncinus typology. ONT sequencing (313 bp COI fragment) significantly outperformed Sanger sequencing, yielding a 78% success rate compared to 36%. Integrated phylogenetic and BOLD database analyses identified six species: four described (*Terebellides shetlandica* Parapar, Moreira & O'Reilly, 2016, *Terebellides europaea* Lavesque, Hutchings, Daffe, Nygren & Londoño-Mesa, 2019, *Terebellides williamsae* Jirkov, 1989, and *Terebellides lavesquei* Barroso, Moreira, Capa, Nygren & Parapar, 2022) and two undescribed lineages (Clade 4 and Clade 12), with *T. shetlandica* dominating relative abundance (80%). *T. shetlandica* and *T. europaea* were the most widespread and northernmost taxa. *T. shetlandica* exhibited the broadest bathymetric range (12–279 m), while Clade 4 was restricted to the shallow Kattegat (12–19 m) and *T. williamsae* to deep offshore waters (179–295 m). Morphological characters were highly congruent with the molecular clades. This study establishes a robust framework for resolving the diversity and distribution of the *Terebellides* species complex, thereby significantly enhancing the accuracy of benthic biodiversity monitoring in Swedish waters.

Introduction

Marine ecosystems

Marine ecosystems constitute an essential component in the functioning of the biosphere by maintaining global biogeochemical cycles (global oxygen, carbon sequestration) (Costanza *et al*., 1997; Halpern *et al*., 2008) and provide numerous ecosystem services to humans, such as climate regulation, coastal protection, food and medicinal products, recreational activities and coastal livelihoods (Peterson & Lubchenco, 1997; Selig *et al*., 2018).

However, marine ecosystems undergo strong anthropogenic pressures such as overfishing, pollution, eutrophication and species introduction (Halpern *et al*., 2008; Hoegh-Guldberg & Bruno, 2010; Burrows *et al*., 2011). This is added by a rapidly changing climate and ocean acidification (Doney *et al*., 2012). These increasing disturbances degrade their functioning and alter their ability to provide goods and services (Worm *et al*., 2006; Crain *et al*., 2008). To address these disturbances, accurate environmental monitoring strategies are required in order to quantify the impact of these multiple stressors. Such insights enable scientists and stakeholders to implement effective countermeasures to mitigate anthropogenic pressures on marine ecosystems.

Environmental monitoring

Environmental monitoring involves gathering samples and measurements of various environmental parameters, such as physical (e.g. temperature, salinity, turbidity), chemical (e.g. pH, nutrient levels, pollutants) and/or biological parameters (Lovett *et al*., 2007). Among biological parameters, biodiversity assessment is central to numerous monitoring programs that track ecosystem changes. Biodiversity, derivative of 'biological diversity' (Wilson, 1985, 1988), is defined as "... the variability among living organisms from all sources including, inter alia, terrestrial, marine, and other aquatic ecosystems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems."(CBD, 1992, p. 4). Evidence suggests that ecosystem functioning—specifically its efficiency and resilience—is enhanced by high levels of biological diversity (Danovaro *et al*., 2008; Palumbi, 2001 ; Palumbi *et al*., 2008 ; Palumbi *et al*, 2009).

Biodiversity can be measured by employing biodiversity indicators to provide reliable estimates of ecosystem status (Noss, 1990; Magurran, 2003). In community ecology, species richness (the number of species present in a community) and relative species abundance (the proportion of each species in a community contributing to the total number of individuals) define the biological structure of an ecological community (Vellend, 2010; Smith & Smith, 2012). An ecological community is defined as an assemblage of populations of different species occupying a same location and interacting with each other directly or indirectly (Briers, 2006; Vellend, 2010; Smith & Smith, 2012). Both species richness and relative abundance allow for the description of a community's species diversity, which can be quantified with help of diversity indices (e.g. Shannon-Weaver index, Simpson index) allowing to characterize the structure of a single community or to compare the structures of multiple communities with one another (Briers, 2006; Smith & Smith, 2012). When considering the spatial organization of ecological communities, beta-diversity—among its multiple versions adapted to specific ecological contexts (Vellend, 2001; Koleff *et al*., 2003; Jost, 2007; Jurasinski *et al*., 2009; Tuomisto, 2010a,b; Socolar *et al*., 2016)—enables the comparison of species diversity between different ecosystem patches or along environmental gradients (Whittaker, 1960, 1972; Anderson *et al*., 2011). While alpha-diversity describes the species richness and relative abundance within a single collection site, and gamma-

diversity represents the total diversity of the entire study area, beta-diversity highlights the variation in species composition between these locations (Whittaker, 1960, 1972; Anderson *et al*., 2011). As such, when moving across an environmental gradient (e.g. bathymetrical gradient in aquatic ecosystems), this spatial indicator allows for the description of a specific type of change in community composition—referred to as species turnover—where species present at one location are replaced by different species at another location that were absent from the previous one (Baselga, 2010).

Consequently, quantifying species diversity and distribution across multiple localities enables the monitoring of how communities shift along environmental gradients and over time.

Marine benthic communities

Marine benthic communities, collectively referred to as the benthos, comprise diverse assemblages of organisms inhabiting the seafloor, ranging from shallow coastal shelves to the deep sea. These organisms reside within or upon varied marine substrates, including sand and both organic and inorganic sediments (Urry *et al*., 2020). Due to their fundamental role in ecological processes, such as nutrient cycling and carbon decomposition, monitoring of marine benthic communities is essential for accurately assessing the ecological status of marine ecosystems (Hauer *et al*., 2018; Kritzer *et al*., 2016). Marine annelids are one of the most dominant taxa among marine communities and play a substantial role in the functioning of benthic communities as they are either predators, herbivores, scavengers or filter feeders (Miller & Harley, 2016; Hutchings, 1998).

Annelids & *Terebellides*

Annelids are one of the most common and abundant taxonomic group encountered within marine habitats and found across a vast array of ecological niches (hydrothermal vents, glaciers, reefs, intertidal zones, etc.) (Rouse, Pleijel & Tilic, 2022). They form a monophyletic group of vermiform invertebrates with metamerism (*meta*, 'after' + *mere*, 'part') as main key character (Miller & Harley, 2016) which refers to a body organization into identical and successive ringlike segments, with many of internal structures repeated within each segment (Miller & Harley, 2016; Urry *et al*., 2020). Annelids are distributed worldwide in marine, freshwater and terrestrial habitats (Struck *et al*., 2014; Miller & Harley, 2016). Based on the relative number of chaetae—bristles made of chitin—on their bodies per segment, this phylum was traditionally subdivided into three main groups: the polychaetes (from the Greek *poly*, 'many', and *khaítē*, 'hair'), the oligochaetes (*olígos*, 'few', and *khaítē*, 'hair') (e.g. earthworms) and achaetes ('hairless'), which refers to Hirudinea (the leeches) (Urry *et al*., 2020). However, these traditional names are no longer used to describe the evolutionary history of the annelids. The term 'polychaete' no longer has taxonomic value as it represents a paraphyletic group; it refers mainly to marine annelids and encompasses the monophyletic group Clitellata (leeches, earthworms, and others) (Miller & Harley, 2016). Recent phylogenomic analyses now divide Annelida into two major derived clades, Errantia and Sedentaria—besides other basal annelid groups (Struck *et al*., 2011; Miller & Harley, 2016)—a grouping that reflects broad differences in lifestyle (Urry *et al*., 2020). Errantia (mostly marine) are primarily active, mobile predators capable of swimming, crawling, or burrowing. In contrast, Sedentaria—which includes marine tubeworms, Clitellata, and other groups—tend to be less mobile; they are often sessile organisms that burrow slowly or live within protective tubes that support their soft bodies. These tube-dwelling sedentarians typically employ filter-feeding or deposit-feeding strategies, facilitated by their elaborate gills or tentacles (Urry *et al*., 2020).

Within Sedentaria, members of the genus *Terebellides* Sars, 1835 (Annelida) are tube-dwelling surface deposit-feeders and are mainly found in soft bottoms on continental shelves and

slopes (Nygren *et al*., 2018). This genus holds the highest species diversity in the Trichobranchidae family with a number of 52 confirmed species recorded (Hutchings *et al*., 2017).

However, it has been shown that annelid identification based on morphological characters alone can seriously underestimate species diversity and that cryptic species are common among these taxa (Nygren, 2014), even though the importance and abundance of this omitted biodiversity has not been sufficiently assessed yet (Nygren, 2014; Bickford *et al*., 2007; Brasier *et al*., 2016).

Species complexes

Like other types of overlooked taxa, species complexes can be a difficult challenge for accurate species monitoring. Species complexes harbor cryptic species that are defined as 'two or more distinct species that are erroneously classified (and thus hidden) under one species name' (Bickford, 2007). The inability of correctly identifying a cryptic species complex can have many negative impacts on species conservation and habitat protection, in addition to many other varied fields (Bickford, 2007). In terms of environmental monitoring, precise identification of cryptic species allows for more accurate results during biodiversity assessment, ecotoxicological analyses and the reading of potential bioindicator species.

Cryptic species might be more common in marine taxa because of several reasons. First, given that distinct morphological features are often correlated with organisms depending on visual signals (Bickford, 2007), marine organisms often rely on chemical cues for mate choice (Stanhope *et al*., 1992) and for gamete recognition (Palumbi, 1994). In addition, there are human reasons for this amount of cryptic species in such environments: the inaccessibility of the marine environment that forces taxonomists to use sample preservation methods which deteriorate or erase certain morphological characteristics necessary for identification (e.g. coloration) and an over-conservative traditional taxonomy that still accepts both too large intra-specific variations and distribution zones of species erroneously considered as cosmopolitan.

Therefore, without using genetic data, it is currently impossible to correctly identify species of *Terebellides* in Sweden.

Taxonomy & DNA barcoding

Taxonomy allows the identification, the classification and the naming of organisms (Winston & Disney, 2000) by using both morphological and molecular data, and based on a pre-established species concept. Given the lack of consensus on a single unanimous species concept and that this issue is not part of this study, we define species here as separately evolving metapopulation lineages, according to de Queiroz's species concept (De Queiroz, 2007).

DNA barcoding is a taxonomic technique that uses the sequencing of a particular DNA fragment—called a barcode—and allows for two distinct processes: species delimitation and species identification. Species delimitation uses multiple molecular markers to infer the boundaries and number of species, especially when those boundaries are unclear within a group of organisms, whereas species identification is the process of assigning an unknown specimen to an already molecularly delimited species by comparing both their morphological and genetic characteristics (De Queiroz, 2007). Our study is based on the use of DNA barcoding as a tool to identify species from collected *Terebellides* specimens. For this purpose, a short sequence (658 nucleotide base pairs) of the mitochondrial gene cytochrome c oxidase subunit I (COI) is widely used as a standardized tool for rapid, robust, and accurate species identification (Hebert *et al*., 2003; Savolainen *et al*., 2005). This mitochondrial gene is a valuable species identification marker in molecular systematics among animals for several reasons. First of all, a mitochondrial marker suits best than a nuclear one because of its lack of introns, limited recombination and its haploid mode

of inheritance (Saccone *et al*. 1999). Then COI is mostly used among other mitochondrial markers because of its conserved sequence (widespread presence among Metazoa), its moderate evolutionary rate and its practical suitability for DNA sequencing (standardized short sequence of DNA facilitating DNA extraction and amplification). Simply put, a DNA barcode library of known species is built, then the barcode sequence of the unknown samples is matched against the barcode library to identify it (Kress & Erickson, 2012). A reference DNA library for several markers, including COI, is available for *Terebellides* species (Nygren *et al*., 2018).

Furthermore, in the present study, we have implemented an enhanced approach of DNA barcoding called DNA mega-barcoding.

DNA mega-barcoding with Oxford Nanopore Technologies (ONT) sequencing

To overcome the limitations of traditional morphology-based identification and the short-read constraints of second-generation sequencers (e.g., Illumina), this study utilizes third-generation sequencing (TGS) for obtaining COI barcodes using MinION nanopore sequencer developed by Oxford Nanopore Technologies (ONT). The MinION (Oxford Nanopore Technologies) is a recently developed nanopore-based DNA sequencing platform and has several advantages over traditional sequencing technologies.

Pooled tagged COI amplicons are obtained from a large number of specimens and are injected into a flow cell set inside the MinION sequencer. The sequencer uses electrical current fluctuations to reconstruct the DNA sequence as the single-stranded DNA molecule translocates through a nanopore (a protein pore of nanometric scale). Sequenced reads are then re-assigned to associated specimen (de-multiplexing) and ONTbarcoder software generates individualized barcode through consensus calling steps (Srivathsan & Meier, 2024).

The main advantage of nanopore sequencing for this study is its high-speed throughput, which allows for large-scale barcoding of thousands of specimens—also known as mega-barcoding (Chua *et al*., 2023). Furthermore, another key advantage of nanopore sequencing—while not essential to the current study—is its ability to generate very long reads (Srivathsan & Meier, 2024). MinION latest improved flow cell now yields highly accurate barcodes that are identical (> 99.99% identity) to aligned barcodes for the same specimens obtained with Sanger and Illumina sequencing (Srivathsan *et al*., 2021). Plus, MinION sequencers are cheap, portable and can generate data onsite and in real time (Srivathsan & Meier, 2024).

Hence, the use of ONT for DNA barcoding seems to be a cost-effective and efficient solution for large-scale specimen identification through DNA barcoding in order to measure species diversity and distribution.

Terebellides in the North East Atlantic

In the North East Atlantic (NEA) waters (ranging from the British Isles in the south, to the Polar Basin in the north), seven species have been described or reported based solely on morphological data (Nygren *et al*., 2018), among which two species in the Swedish coastlines (*Terebellides stroemii* & *Terebellides gracilis*) (SLU Artdatabanken 2025). After sequencing and analyzing several DNA markers from over 500 specimens, a hidden species complex within the genus *Terebellides* was discovered (Nygren *et al*., 2018). A total of 27 species were detected in the NEA waters instead of the seven previously known, most of which were described (Nygren *et al*., 2018). From these 27 species, ten are present in the Skagerrak and Kattegat, of which only eight are described.

Aim

Our study has three main objectives. The first is to identify all *Terebellides* specimens collected by the Swedish national environmental monitoring program (PMK) in 2022 and 2024 by applying a dual-platforming approach—integrating both Sanger and ONT sequencing—to assess species diversity and distribution. The second aim involves comparing the performance of these two sequencing platforms throughout the study. Finally, the third aim is to identify diagnostic morphological characters for any undescribed *Terebellides* species revealed by the molecular data.

Material and methods

Study area & sampling

Specimens were collected from marine bottom sediment sampled in 2022 and 2024 as part of the Swedish national environmental monitoring program PMK (*Program för miljökvalitetsövervakning*). The samples originate from a total of 116 collecting sites along the Swedish western continental coast (65 sites in Skagerrak and 51 sites in Kattegat) between depths of 12 m to 295 m. Skagerrak and Kattegat are shallow straits with Skagerrak having a greater average depth, as it is connected to the Norwegian trench. Sediment distribution in both straits includes deep bottoms covered with clays and fine silts, while shallow coastal areas are dominated by mud, sand and stones (Svansson, 1975; Bengtsson & Stevens, 1996). Sediment samples of 16.6L were taken using a Smith-McIntyre grab. Samples were sieved on a 1 mm sieve on board, bulk fixed in 95% ethanol with 5% glycerol. Samples were then carefully sorted and morphologically identified by an expert to genus level (*Terebellides*) and stored at Tjärnö marine laboratory. Among these *Terebellides*, a total of 529 specimens were selected, named and labeled in a database for molecular and morphological analysis. Sampling did not include endangered or protected species.

DNA extraction, amplification & sequencing

DNA extraction

A tissue sample was taken from each selected individual. A lateral piece of tissue between the thorax and abdomen was taken for large specimens, and the entire abdomen up to the pygidium for very small specimens. Tissue sampling was performed using tweezers, a petri dish filled with ethanol and a stereo microscope (Leica M165C stereomicroscope). The head was always preserved for later morphological observation. The tissues were placed and preserved in labeled eppendorf tubes. DNA extraction was done using QuickExtract following manufacturer's protocol. This solution contains a mixture of enzymes, including proteinase K. This enzyme lyses cell walls and other cellular components, thus extracting DNA molecules from the cells. To do this, 50 μL of this solution were added to each individual tissue. The samples were then mixed and heated at 65°C for 2 hours using a thermal stirrer. After two hours, the samples were heated at 95°C for 5 minutes to degrade proteinase K, which risks to alter the polymerase enzyme used in the next step of the PCR process. Once this step was complete, the samples were stored in the freezer.

DNA amplification

PCR was performed twice for each sample. Once to obtain amplified DNA for Sanger sequencing, and once to obtain DNA for sequencing with ONT.

First, a Master Mix was prepared containing Taq DNA polymerase and a pair of primers. As target primers for the standard mitochondrial gene COI (cytochrome c oxidase subunit I), several nucleotide primers were tested upstream, and mlCOIintF (forward) and COIE (reverse) were chosen for Sanger sequencing. Several PCR plates were filled with this Master Mix solution (10.6 μL for each well), to which 1.4 μL of extracted DNA from each specimen was added individually. Thus, 0.6 μL of primers (0.3 μL forward and 0.3 μL reverse); 10 μL Taq DNA Polymerase and 1.4 μL DNA were used for each individual sample. The plates were then sealed with caps and centrifuged. For Sanger sequencing, the same primers were used for all samples. For sequencing with ONT, six different forward primers were used consisting of the mlCOIintF primer with a 9-bp tag attached to it and 95 different reverse primers were used consisting of the COIE primer with a 9-bp tag attached to it. This allowed to double-indexing each sample with a unique combination of forward and reverse tags.

To multiply the target sequence in large quantities for each of our specimens, the PCR plates containing the samples were placed in a thermal cycler, which allows the polymerase chain reaction technique (PCR) to be performed. As with the primers, several tests with different programs were performed beforehand to select the optimal program. The samples underwent several cycles of a three-step chain reaction. The cycle consisted of an initial denaturation step at 95°C for five minutes, followed by five cycles composed of a denaturation step at 95°C for 40 seconds, an annealing step at 45°C for one minute and an elongation step at 72°C for one minute, followed by 35 cycles as preceding but with an annealing temperature of 48°C, followed by a final elongation step at 72°C for seven minutes.

To verify the quality and quantity of the PCR-amplified DNA samples for Sanger sequencing, agarose gel electrophoresis was performed. No check of the PCR products is required for sequencing with ONT, so these samples were not tested on agarose gel. 1.5 g of agarose powder was weighed and poured into a glass jar into which 100 mL of 1x TAE buffer solution was added. This mixture was then heated in the microwave for 1 minute 30 seconds (at medium temperature). Once ready, 5 μL of fluorescent dye (GelRed) was added. The mixture was then poured into a gel casting tray after placing the comb into the tank. The gel solidified after 20-30 minutes. Once the gel was ready, the comb was gently removed, leaving wells in the gel, and the casting tray containing the gel was then placed inside the electrophoresis tank filled with TAE buffer solution. 2 μL of each PCR sample were inserted separately into an individual well of the gel. A 100-volt electric field was then applied for 25 minutes to allow the samples to migrate. After migration, the gel was placed inside the dark chamber of a transilluminator (Axygen). Once the gel was illuminated with UV light, fluorescent bands reflecting the quality and quantity of the PCR products could be observed.

DNA sequencing

One set of high-quality PCR products were sent at Eurofins Genomics (Germany) for purification and Sanger sequencing on both strands after being added individually to 10 μL of water in new adapted PCR plates (96-well PlateSeq Kit PCR). Following sequencing, forward and reverse reads were assembled into consensus sequencing, trimmed from primers and quality checked in Geneious 5.4.6. (https://www.geneious.com). Each sequence was checked for contamination using BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and non-annelid sequences were removed.

Following amplification, the other set of PCR products for sequencing with ONT were pooled and 1mL of the pooled products was purified using a 1x ratio of AMPure XP beads

(Beckman Coulter) following the protocol described in Srivathsan and Meier (2024). The DNA library was prepared using ONT ligation kit V14 (SQK-LSK114) and following the manufacturer's protocol for amplicon sequencing, with the exception that all purification steps were performed with a 1x ratio of AMPure XP beads as recommended by Srivathsan and Meier (2024) for short amplicons. The library was sequenced on a flow cell with R.10.4.1 chemistry on a MinION Mk1C device running MinKNOW 24.11.8. Fast accuracy basecalling was performed directly on the MinION Mk1C. Reads were then demultiplexed, trimmed, assembled and automatically quality checked in ONTbarcoder 2.3.0, which is a software dedicated to fully and automatically process this type of data from end to end. Default parameters were used with a target length of 313 bp and a minimum read length of 100 bp. The dataset was checked for contamination using BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and non-annelid sequences were removed. All the QC compliant barcodes of full length and with no ambiguities were used in downstream applications.

Sequence & distribution analyses

The barcodes were analyzed from two separate fasta files, one containing the Sanger sequences and the other the ONT sequences. All barcodes contained in these files are consensus sequences that have been filtered, cleaned and trimmed beforehand. The barcodes obtained by the two methods were all grouped together, taking care to remove duplicates belonging to the same specimen. To these barcodes were then added 82 homologous sequences, which act as reference sequences. These reference sequences belong to clades that have already been genetically established in Nygren *et al*. (2018) by means of several species delimitation analysis (GMYC, TCS and STACEY) based on both mitochondrial (COI, 16s rDNA) and nuclear (ITS2, 28s rDNA) genetic markers. Only reference sequences belonging to clades relevant to our study were selected, obtaining a total of 82 sequences from six already established species (Clades 1, 2, 4, 5, 6 and 12). All sequences were imported in AliView (version 1.30) (Larsson, 2014) and then aligned using the MUSCLE alignment tool (Edgar, 2004) available in AliView. The output file in fasta format was then imported into the Galaxy online platform (The Galaxy Community, 2024) for phylogenetic inference. IQTREE (Minh *et al*., 2020) was used to generate a phylogenetic tree using the maximum likelihood probabilistic method (ML) while setting the bootstrap to 1000 replicates (Hoang *et al*., 2018). According to the Akaike Information Criterion (AIC), the best-fitting evolution models (GTR+F+I+R4) used by the ML method were automatically selected by ModelFinder (Kalyaanamoorthy *et al*., 2017) incorporated in IQTREE. Once the phylogenetic reconstruction was complete, the NEWICK code produced by IQTREE was copied and then imported into the web-visualizer iTOL (v6) (Letunic & Bork, 2024) so that the phylogenetic tree could be viewed and annotated. Thus, each sequence to be identified was grouped with reference sequences (sequences of already known species) with which it shares the least phylogenetic distance, thereby revealing the species to which the specimens to be identified belong. In addition, many sequences were also blasted to the BOLD (Barcode of Life Data System) reference database (Ratnasingham & Hebert, 2007) to ensure the accuracy of the phylogenetic tree results.

Geographic distribution maps (Figure 2) were generated using QGIS (version 3.40.5-Bratislava) (QGIS.org, 2026) based on the longitude and latitude data of each specimen. The dot-plot illustrating the bathymetric distribution of the species (Figure 3) was generated using Microsoft Excel (version 16.16.2) (Microsoft Corporation, 2018) based on the bathymetric data of each specimen.

Morphological analyses

Morphological analyses were performed last and focused solely on specimens belonging to Clades 4 and 12, as these two species are undescribed. They focused on three criteria: branchial shape, uncini morphology (thoracic and abdominal) and methylene blue staining patterns. Other remaining criteria require a more thorough and precise examination, necessitating the use of a scanning electron microscope (SEM) which was not available here.

First, the observation of branchial shape was carried out using a binocular microscope (Leica M165C stereomicroscope) and the designation of the corresponding type is based on the ones defined in Parapar *et al*. (2016). Branchial shape examination was carried out on four specimens belonging to Clade 12 (PA115A; PA87A; PA85A; PA7A) and nine specimens belonging to Clade 4 (PA25A; PA32B; PA32C; PA91C; PA92B; PA92C; PA92D; PA92E; PA134D).

Next, the identification of thoracic and abdominal uncini types was based on dissected parapodia that were mounted on slide to be observed with an Olympus BX51 microscope. The parapodia were taken from two specimens, one from each clade (PA87A for Clade 12 and PA134D for Clade 4). Uncini are specific types of chaetes that anchor the polychaete to the substrate thanks to their hook-like shape and are present from the sixth chaetigerous segment onward (Rouse, Pleijel & Tilic, 2022). Designation of thoracic and abdominal uncini types is based on the types defined respectively in Parapar *et al*. (2020b) and Parapar *et al*. (2020a).

Finally, methylene blue staining involved immersing a specimen in one solution of methylene blue before rinsing it with ethanol. The specimen is then placed in a Petri dish filled with ethanol and observed under a stereo microscope. The usefulness of this technique lies in the stain's ability to adhere unevenly to the surface of specific thoracic segments, resulting in the formation of pigment patterns characteristic of the species to which the dyed specimen belongs. Considering that methylene blue (MB) and methyl green (MG) share similar results, the identification of the type of patterns obtained was based on the description and definition of MG staining patterns in Schüller & Hutchings (2010) and Schüller & Hutchings (2013). MB staining was carried out on four specimens belonging to Clade 12 (PA115A; PA87A; PA85A; PA7A) and nine specimens belonging to Clade 4 (PA25A; PA32B; PA32C; PA91C; PA92B; PA92C; PA92D; PA92E; PA134D).

Results

Sequences obtained by both sequencing methods

A total of 529 specimens were selected for sequencing, each specimen having been sequenced twice (both with Sanger and ONT methods). The exact sequence length used for alignment was 313 bp, corresponding to a mini-barcode known as the 'Leray region' (Leray *et al*., 2013). This region is embedded within the standard 658-bp 'Folmer fragment' (Folmer *et al*., 1994) of the COI gene. This short segment is sufficiently long and informative to allow species identification in this study. Of the 529 specimens examined, 413 ONT barcodes were called, giving a barcoding success rate of 78%, compared to 191 Sanger barcodes successfully called (36%). No ONT sequences showed any gaps, whereas three Sanger sequences were incomplete: PA91C (Clade 4; last 18 nucleotides missing), PA92D (Clade 4; last 18 nucleotides missing), and PA19A (*T. shetlandica*; first 15 nucleotides missing). Specimens that were successfully sequenced by both methods never showed dissimilar sequences that could hinder species identification. Of the 191 successful Sanger

barcodes, 17 were not obtained by ONT (PA54B; PA128C; PA24B; PA111A; PA94D; PA146E; PA117E; PA135D; PA99A; PA124E; PA144A; PA115A; PA92D; PA118C; PA52E; PA1C; PA148A). The 413 ONT sequences were added to the remaining 17 Sanger sequences to obtain a total of 430 barcodes to align with reference sequences for phylogenetic analyses (Figure 1).

Species richness & relative abundance

The analyses of the 430 barcodes to be determined, based on a phylogenetic reconstruction integrating additional reference sequences, as well as on the blasting of these barcodes to the BOLD reference database (Ratnasingham & Hebert, 2007), made it possible to highlight the presence of six species belonging to the genus *Terebellides* (*T. shetlandica*, *T. williamsae*, Clade 4, *T. lavesquei*, *T. europaea* and Clade 12) along Sweden's west coast (Figure 1).

In terms of relative abundance, the most frequently encountered species was *T. shetlandica* with 343 specimens (80% of the sequences). Next, in descending order, were *T. europaea* with 39 specimens (0.09%), Clade 12 with 18 specimens (0.04%), Clade 4 with 13 specimens (0.03%), *T. williamsae* with 7 specimens (0.02%), and finally *T. lavesquei* with 3 specimens (0.007%).

Among these 430 sequences, seven sequences came from individuals not belonging to the genus *Terebellides*: *Melinna palmata* (PA24B; PA129A; PA147A), *Melinna cristata* (PA143C), *Lysilla loveni* (PA33D; PA33E) and *Pista lornensis* (PA62B) (Figure 1).

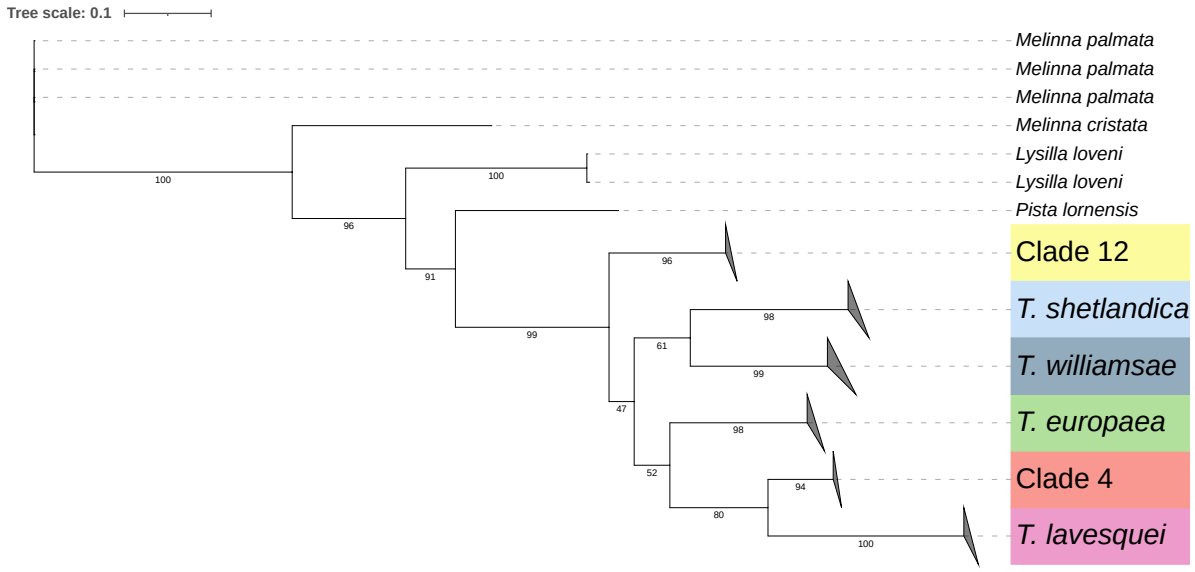


Figure 1. Phylogenetic relationships based on mitochondrial COI mini-barcode (313bp) of all studied specimens (430 sequences) added to reference sequences (82 sequences). The phylogenetic tree was reconstructed using the Maximum Likelihood (ML) method with IQ-TREE on the Galaxy platform. [GTR+F+I+R4] model was selected by ModelFinder. Numbers below branches indicate branch support assessed using 1000 bootstrap replicates. The tree was rooted on the outgroup consisting of three identical sequences belonging to *Melinna palmata*. The scale bar represents the average number of nucleotide substitutions per site. Colored leaves correspond to collapsed clades containing specimens belonging to identified *Terebellides* species (6 species). Non-colored leaves correspond to sequenced specimens not belonging to *Terebellides* genus (7 individuals).

Geographic & bathymetric distributions

Of the six identified *Terebellides* species, *T. shetlandica* and *T. europaea* showed the widest distribution ranges (Figure 2). Furthermore, the specimens collected in the northernmost stations of the study area also belong to these two same species (Figure 2). Besides, Clade 12 also showed a fairly wide distribution range compared to Clade 4, which seems to be limited to the Kattegat (Figure 2). Finally, the three collected specimens belonging to *T. lavesquei* all come from the same place, the Gullmarsfjorden, connected to the Skagerrak, while the seven specimens belonging to *T. williamsae* were collected in the Skagerrak, but solely away from the coastline (Figure 2).

The two most widespread species in the study area, *T. shetlandica* and *T. europaea*, showed similar bathymetric distributions with a majority of specimens having been collected between 12 and 110 m deep (Figure 3). However, *T. shetlandica* was also found beyond 250 m depth while *T. europaea* did not exceed 100 m. Another fairly well-distributed species, Clade 12, similarly to *T. europaea*, did not exceed 100 m in depth (Figure 3). The three specimens of *T. lavesquei*, all from the same area, were all collected at approximately the same depth (100 m) (Figure 3). No *T. williamsae* specimens have been collected before 179 m depth and they showed the highest bathymetric value (up to 295 m) (Figure 3). Finally, specimens belonging to Clade 4 were not collected beyond a depth of 20 m (Figure 3).

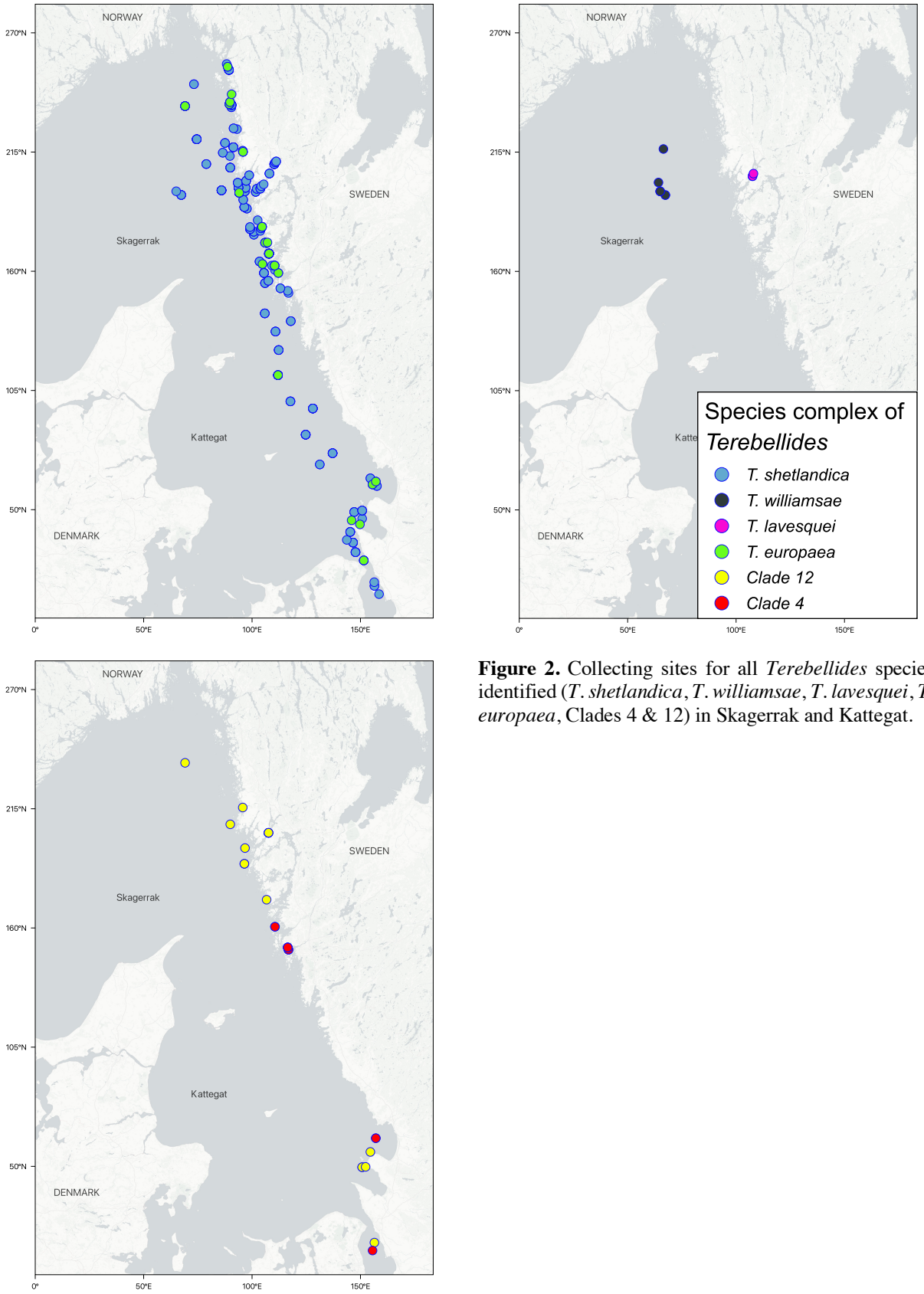


Figure 2. Collecting sites for all *Terebellides* species identified (*T. shetlandica*, *T. williamsae*, *T. lavesquei*, *T. europaea*, Clades 4 & 12) in Skagerrak and Kattegat.

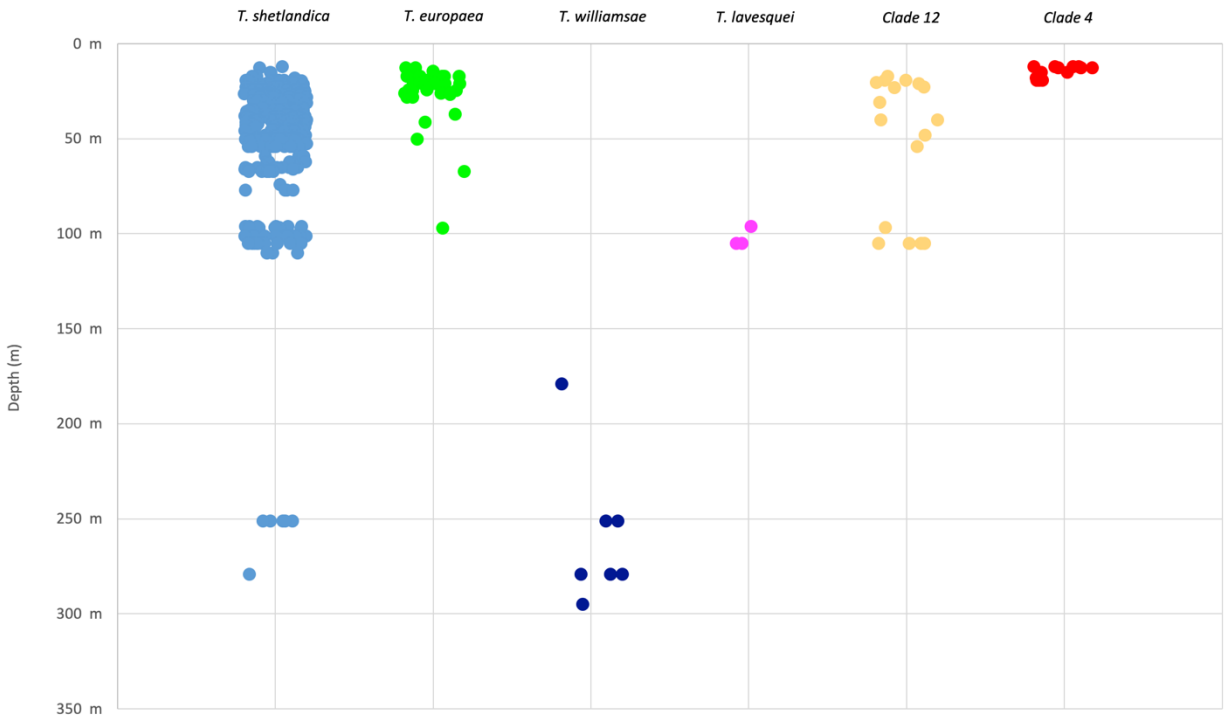


Figure 3. Bathymetric distribution map for *T. shetlandica*, *T. williamsae*, *T. lavesquei*, *T. europaea*, Clades 4 & 12 in Skagerrak and Kattegat.

Morphological analyses

As previously stated, morphological analyses only concerned undescribed species (Clade 4 and Clade 12).

Firstly, branchial shape of Clade 12 appears to correspond to type 1 (fairly fused branchial lobes, a large number of packed lamellae in the dorsal lobes and overall comma-shaped branchiae). Meanwhile, branchial shape of Clade 4 seems to correspond more closely to type 3. Indeed, although the ventral lobes are less prominent than the dorsal ones, there are no comma-shaped branchiae, and the lobes indicate a lesser degree of fusion. Furthermore, the anterior extension of the dorsal lobes (fifth lobe) appears to be absent.

Next, in Clade 12, the thoracic uncini possess an imposing rostrum (RvC 2:1) with a capitium bearing medium-length teeth surmounted by slightly smaller teeth, which would correspond to type 3. Regarding the abdominal uncini, the lengths of the rostrum and capitium are quite similar (RvC ~ 1:1) with a complex capitium composed of a first row of teeth surmounted by a multitude of smaller teeth decreasing in size and number, which would correspond to type 2. In Clade 4, the thoracic uncini also possess a large rostrum (RvC 2:1) with a capitium containing medium-length teeth surmounted by slightly smaller teeth, again suggesting type 3. And regarding abdominal uncini for Clade 4, an equivalent RvC added to a complex capitium also indicate abdominal uncini of type 2.

Finally, the methylene blue (MB) staining profile of Clade 12 indicates a solid staining of the chaetigerous segments TC1 to TC6, which then transitions to a striated staining from TC7 to TC12 before fading after segment TC12. In addition, a J-shaped glandular region is noticeable on segment TC3. The MB staining profile therefore appears to correspond to pattern 1. Regarding Clade 4, MB staining similarly indicates solid staining of the chaetigerous segments TC1 to TC6,

which then evolves into a striated staining from TC7 to TC12 before fading after segment TC12. However, a more pronounced and repeated glandular region is observed across several segments (TC1-TC3). Thus, the sum of these characteristics would approximate pattern 9.

Discussion

Comparative performances between both sequencing methods

A central methodological focus of this study was to conduct a comparative performance analysis between traditional Sanger sequencing and Oxford Nanopore Technologies (ONT). This comparison evaluated the efficacy of both platforms to generate reliable COI barcodes, focusing on their capacity to accurately characterize species diversity and distribution across a high-throughput dataset of *Terebellides* specimens.

The evaluation of sequence quality demonstrated the superiority of ONT, which consistently generated full-length 313 bp reads. In contrast, Sanger sequencing produced several sequences truncated at the extremities. These missing nucleotides at the extremities are generally explained by an overly intense fluorescent signal at the beginning of reads—due to primer dimers or excessive dye-labeled nucleotides leftovers (dyeblob) causing low-quality sequences (Draper, 2008; Crossley *et al*., 2020; Al-Shuhaib & Hashim, 2023)—while fluorescent signal decays towards the ends of reads due to diffused fragment traces leading to reduced base-calling accuracy (Ewing *et al*., 1998; Ewing & Green, 1998). Unlike Sanger, ONT does not rely on the synthesis of new fragments but rather on the direct reading of the DNA strand from start to finish through a nanopore with constant signal intensity, facilitated by ligation adapters (Srivathsan & Meier, 2024). This fundamental difference explains why the ONT sequences remained complete without any gaps. This technical reliability directly translates into a higher quantity of identified specimens; ONT achieved a 78% success rate (413 barcodes) compared to only 36% for Sanger, proving its superiority for high-throughput processing. This technology (ONT) allows for the parallel processing of hundreds of samples in a single run through multiplexing (Srivathsan *et al*., 2019; Srivathsan *et al*., 2021; Srivathsan & Meier, 2024), whereas Sanger sequencing remains constrained by individual sample processing and a labor-intensive workflow necessitating, for example, complex equipment maintenance and dedicated support staff (Shendure & Ji, 2008). Beyond raw performance, the sequencing cost becomes significantly more advantageous with ONT, being highly cost-effective for sequencing large sample sizes (Srivathsan *et al*., 2021). In contrast, Sanger sequencing costs are incremental with a fixed price per sample, exclusive of shipping and external service fees. For instance, the commercial pricing at Eurofins Genomics for obtaining 529 barcodes from already amplified samples exceeded €4 per barcode, based on the use of six 96-well PlateSeq Kit PCR units (Eurofins Genomics, 2025). Once the initial investment in equipment (MinION Mk1C sequencer) has been made, the cost for sequencing consumables—including the R10.4.1 Flow Cell, Ligation Kit (SQK-LSK114), Native Barcoding Kit (SQK-NBD114), and AMPure XP Beads—reaches approximately €4 per barcode for the 529 specimens analyzed in this study. When excluding shared laboratory costs (upstream DNA extraction and amplification reagents and consumables), the ONT approach utilized in this study is slightly more cost-effective than traditional Sanger sequencing services. However, had the number of specimens been doubled, for instance, the cost of sequencing consumables (Flow Cell and kits) would have remained the same, driving the price down to approximately €2 per barcode. Furthermore, Srivathsan *et al*. (2021) report a cost of less than $1 per barcode for large-scale projects that exploit the maximum capacity of the Flow Cell (approximately 10,000 specimens). This cost-effectiveness

is coupled with a significant gain in time; excluding DNA extraction, our ONT workflow provided results in less than 24 hours, compared to the three-day to one-week delay time for Sanger sequencing services. Furthermore, the manual tasks of trimming, filtering, editing, and exporting inherent to Sanger sequencing are replaced in the ONT workflow by an integrated automated pipeline; specifically, ONTbarcoder performs read filtering and trimming before generating the final sequence through consensus calling. Finally, the practicality and democratization of this technology are ensured by the portability of the MinION and the accessibility of the free ONTbarcoder software, offering total autonomy by eliminating reliance on centralized and expensive sequencing platforms (Pomerantz *et al*., 2018; Chang *et al*., 2020; Watsa *et al*., 2020; Pomerantz *et al*., 2022; Srivathsan & Meier, 2024).

In conclusion, the ONT workflow establishes itself as a robust, rapid, autonomous, and economical method for modernizing the genetic identification of species complexes.

Species richness, relative abundance & distribution

As previously stated, the identification of *Terebellides* species in this study is based on the molecular-based delimitation and phylogenetic analyses performed by Nygren *et al*. (2018). Their work distinguished a set of 27 clades subdivided into four distinct phylogenetic groups (Groups A, B, C and D) which together form the *Terebellides* species complex across the North East Atlantic (NEA). Prior to the establishment of these 27 molecular clades, only seven species had been described in this area: *T. stroemii* Sars, 1835; *T. gracilis* Malm, 1874; *Terebellides atlantis* Williams, 1984; *T. williamsae* Jirkov, 1989; *Terebellides irinae* Gagaev, 2009; *Terebellides bigeniculatus* Parapar, Moreira & Helgason, 2011; and *T. shetlandica* Parapar, Moreira & O'Reilly, 2016. Subsequently, several of these clades were the subject of formal taxonomic descriptions. Within just four years, the number of described species in the NEA increased from seven to 14, following the descriptions of: *T. europaea* Lavesque *et al*., 2019; *Terebellides bakkeni* Parapar *et al*., 2020; *Terebellides kongsrudi* Parapar *et al*., 2020; *Terebellides ronningae* Parapar *et al*., 2020; *Terebellides norvegica* Parapar *et al*., 2020; *Terebellides scotica* Parapar *et al*., 2020; and *T. lavesquei* Barroso *et al*., 2022. Several clarifications have also been made. Notably, Lavesque *et al*. (2019) molecularly assigned Clade 6 to *T. europaea*; this clade had been erroneously identified as *T. stroemii* by Nygren *et al*. (2018). Furthermore, a distinction was established between the morphologically similar *T. williamsae* (Clade 2) and *T. gracilis* (Clade 3). Although Parapar *et al*. (2011) originally suggested that the former was a junior synonym of the latter based on Icelandic specimens, Nygren *et al*. (2018) demonstrated their molecular distinction using Barents Sea specimens. Finally, Barroso *et al*. (2022) showed that *T. williamsae* differs morphologically from *T. gracilis* in the distribution pattern of white ventral thoracic chaetigers coloration (TC1–4 in *T. williamsae* and only on TC4 in *T. gracilis*). Consequently, of the 27 clades detected in the NEA, 12 remain undescribed: Clades 4, 12, 14, 15, 17, 18, 19, 22, 23, 25, 26, and 27. Within the Kattegat and Skagerrak—a smaller subset of the NEA—ten clades have been detected. Currently, 8 of these are formally described (*T. shetlandica*, *T. williamsae*, *T. gracilis*, *T. lavesquei*, *T. europaea*, *T. ronningae*, *T. norvegica* and *T. kongsrudi*) while two remain undescribed (Clades 4 and 12). In the present study, a total of six *Terebellides* species were identified within the Kattegat and Skagerrak regions. Of these, four are described taxa (*T. shetlandica*, *T. williamsae*, *T. lavesquei* and *T. europaea*) while the remaining two match with Clades 4 and 12.

In this study, *T. shetlandica* was found to be the most widely distributed species across the study area, as well as the most abundant (343 specimens, contributing to 80% of the sequences). It was recorded from the northern Skagerrak to the southern Kattegat, primarily at depths between 12 and 110 m, though it was occasionally found deeper (251–279 m). These findings are consistent with previous observations in the Kattegat (approximately 30 m) and the Skagerrak (up to 250 m) (Nygren *et al*., 2018). Across the NEA, this species is also present along the Norwegian coast and

shelf (30–375 m) and in the North Sea (76–246 m), with the majority of specimens collected below 200 m (Nygren *et al*., 2018). This species corresponds to Clade 1 according to Nygren *et al*. (2018) and belongs to Group B. Its type locality is situated between the Shetland Islands and the Norwegian coast at a depth of 160 m. Consequently, *T. shetlandica* appears to be a common species ranging from the Shetlands across the Norwegian shelf to the Öresund strait, occupying a broad bathymetric range. *T. shetlandica* manifests therefore as a ubiquitous species that is seemingly unrestricted by specific habitat types.

Next, *T. europaea* was the second most frequently identified species (39 specimens, contributing to 9% of the sequences) and possesses a wide geographic distribution similar to *T. shetlandica*, ranging from the northern Skagerrak to the Öresund in the south. However, unlike *T. shetlandica*, it was not recorded at depths exceeding 100 m. This is consistent with previously recorded bathymetric data for this species, which indicate a depth of 38 m in the Kattegat and no occurrences beyond 100 m in the Skagerrak (Nygren *et al*., 2018). Outside the Kattegat and Skagerrak within the NEA, this species has also been found at shallow depths along the Norwegian coast and shelf (13 m), in the North Sea (85 m) and in the Irish Sea (42–54 m) (Nygren *et al*., 2018). This species corresponds to Clade 6 (Group A) according to Nygren *et al*. (2018) and its type locality is the Bay of Brest (Brittany, France) (Lavesque *et al*., 2019). While displaying a broad geographic distribution, this species exhibits a distinct preference for shallow coastal habitats. Furthermore, its distribution extends southward beyond the NEA, reaching the Bay of Biscay and as far as the Ría de Ferrol in Galicia, Spain (Lavesque *et al*., 2019).

In this study, *T. williamsae* was recorded only in the Skagerrak and exclusively at depths exceeding 179 m. While this restricted observation could be attributed to the small sample size (7 specimens, representing approximately 1.6% of the sequences), previous research confirms that nearly all specimens of this species have consistently been collected at depths greater than 200 m (Nygren *et al*., 2018). Regarding the specific study area, *T. williamsae* has previously been identified in the Skagerrak (between 224–335 m) but has never been recorded in the Kattegat, which is a shallow strait. The species has also been found along the Norwegian coast and shelf (357–534 m), the Norwegian Sea (610–612 m), the Barents Sea (178–378 m) and further north in the Greenland Sea (315–317 m). *T. williamsae* is assigned to Clade 2 within Group D and its type locality is the Barents Sea, located between northern Norway and Svalbard, at a depth of 385–390 m (Nygren *et al*., 2018). Consequently, this species appears to favor deep-water habitats and exhibits a more northern distribution range compared to the two aforementioned species.

Next, the last described species identified in this study was *T. lavesquei*, for which all specimens were collected at a single location (Gullmarsfjorden) in the Skagerrak, at a depth of approximately 100 m. It is highly probable that the small sample size (3 specimens, contributing roughly to 0.7% of the sequences) is not representative of the species' full distribution, unless it is a rare species restricted to a very localized area. However, previous research reports two specimens collected at approximately 250 m in the Skagerrak and 17 specimens between 115 and 534 m along the Norwegian coast and shelf. *T. lavesquei* is assigned to Clade 5 and belongs to Group B (Nygren *et al*., 2018). While exhibiting a relatively broad bathymetric range, the distribution area of this species appears limited, with nearly all specimens collected exclusively along the Norwegian coast and shelf. Regarding the study area, *T. lavesquei* seems relatively rare in the Skagerrak and is absent from the Kattegat. The collection of additional specimens may provide further clarity regarding its distribution within the studied region.

Several specimens belonging to two as-yet undescribed species were identified in this study (Clades 4 and 12). These species are currently being described elsewhere by D. Gaeva and I. Jirkov (Shirshov Institute of Oceanology, Russia). The present study focuses exclusively on observing their distribution within the Skagerrak and the Kattegat. Among these two undescribed species, most specimens belong to Clade 12 (18 specimens, representing approximately 4% of the sequences). The geographic distribution of Clade 12 within the study area is relatively broad,

extending from the Skagerrak to the Öresund in the southern Kattegat. Furthermore, it is notable that no specimens were collected at depths exceeding 100 m. This aligns with previously recorded distributions in the Kattegat (21–22 m) and the Skagerrak (~100 m). However, this clade has been documented at greater depths in the Barents Sea (55–171 m; Nygren *et al*., 2018) and the Kandalaksha Bay of the White Sea (10–170 m; Kudryavtseva *et al*., 2023); nevertheless, its distribution does not appear to exceed a depth of 200 m. According to Nygren *et al*. (2018), Clade 12 belongs to Group A. Consequently, this species appears to be quite frequent along the Swedish West Coast but shows a distinct preference for shallow waters.

In contrast to Clade 12, Clade 4 was found exclusively in the Kattegat. All 13 specimens were recovered from depths of 20 m or less, representing the shallowest mean bathymetric distribution recorded in this study. This appears to align with previous observations, where this species had been recorded at very shallow depths in both the Skagerrak (13.5 m) and the Kattegat (18–20 m) (Nygren *et al*., 2018), but also more recently in Kandalaksha Bay of the White Sea (10-25 m) (Kudryavtseva *et al*., 2023). Although the current sample size may seem insufficient for definitive conclusions, the data from the 13 specimens collected in this study—supplemented by the 18 specimens from the two previous studies—strongly suggest that this clade is restricted to shallow-water coastal habitats.

Finally, all six species identified in this study had been previously reported within the study area (Kattegat and Skagerrak) (Nygren *et al*., 2018). However, some species previously recorded in this region were absent from our study. These four missing species are already described species: *T. gracilis* (Clade 3), *T. ronningae* (Clade 7), *T. norvegica* (Clade 8) and *T. kongsrudi* (Clade 13) (Nygren *et al*., 2018). *T. gracilis* has been recorded in the Skagerrak at depths between 237 and 406 m and lives in sympatry with *T. williamsae*, which was formerly considered the same species. Outside the study area, these two species share similar biogeographic regions (Norwegian coast and shelf, Norwegian Sea) with overlapping bathymetric ranges. Nevertheless, *T. gracilis* occupies greater depths (up to 1268 m on the Norwegian coast and shelf) and possesses a more restricted distribution range. It is probable that a higher sampling effort in the Skagerrak at depths exceeding 200 m would have allowed for the detection of *T. gracilis* individuals during this study. Furthermore, the three other species not recovered in this study previously exhibited varied distributions. *T. ronningae* possesses the most restricted distribution (Skagerrak, Norwegian coast shelf) and the shallowest bathymetry (62–188 m); *T. norvegica* exhibits the deepest bathymetry (209–1268 m); and *T. kongsrudi* occupies the broadest distribution range (Skagerrak, Norwegian coast shelf, Barents Sea, and Greenland Sea) across intermediate depths (126–477 m) (Nygren *et al*., 2018). Within the Skagerrak, the number of specimens previously collected for *T. ronningae*, *T. norvegica*, and *T. kongsrudi* was low (2, 9, and 4 individuals, respectively). Continued sampling will eventually increase the probability of occurrence for these species, allowing for a better assessment of their distribution within the study area.

In summary, co-occurrence patterns within these two regions reflect distinct strategies of habitat occupancy. Certain species exhibit restricted bathymetric ranges—such as Clade 4 in shallow waters and *T. williamsae* in the deep sea—resulting in non-overlapping distributions that prevents physical encounter. Conversely, other species occupy broader ranges, leading to sympatry both among themselves (e.g., *T. shetlandica* and *T. europaea*) and overlapping with the more range-restricted species.

Morphological analyses on Clades 4 & 12

Morphological analyses revealed similar recurring characteristics between specimens belonging to the same clade.

Firstly, according to the description of branchial shape types established in Parapar *et al*. (2016), branchial shape of Clade 12 appears to correspond to type 1. This would be similar to the type already found for this clade in Parapar *et al*. (2020a) (fairly fused branchial lobes, a large number of packed lamellae in the dorsal lobes and overall comma-shaped branchiae). Meanwhile, the branchial shape of Clade 4 seems to correspond at first sight to type 3, the same as *T. shetlandica* and *T. atlantis* from the same phylogenetic group (Group B) (Nygren *et al*., 2018). However, ventral lobes look still less prominent than dorsal ones, there are no comma-shaped branchiae, a lesser degree of fusion of the lobes and the fifth lobe seems absent.

Next, the identification of the morphological type of thoracic and abdominal uncini follows the ones established respectively in Parapar *et al*. (2020b) and Parapar *et al*. (2020a). The uncini typology is based on the rostrum vs capitium length ratio (RvC) as well as on the relative size and number of capitium teeth. The thoracic and abdominal uncini observed on Clade 12 slides remain quite similar to the types already proposed in Parapar *et al*. (2020a) where this clade had already been the subject of a very brief description. Meanwhile, the slides of the thoracic and abdominal uncini of Clade 4 show a morphology similar to that of Clade 12. Clade 4 uncini morphology seems to fit with genetically close *T. lavesquei* and *T. atlantis* which belong to the same clade (Group B) molecularly circumscribed in Nygren *et al*. (2018) and morphologically described in Barroso *et al*. (2022).

In Clade 12, the MB staining profile corresponding to pattern 1, agrees with the pattern previously observed in Parapar *et al*. (2020a), apart from the recurrent glandular region on TC3 which is not mentioned in this paper. In Clade 4, the MB staining profile seems to approximate pattern 9. Again, this would be consistent with the same pattern as that of the genetically close species of group B (*T. lavesquei* and *T. atlantis*) described in Barroso *et al*. (2022).

Despite some compelling observations, the morphological analyses carried out here on the two undescribed species (Clades 4 and 12) remain incomplete due to a lack of observed criteria (papillae on branchial lamellae margin and ciliated papillae dorsal to thoracic notopodium) due to the absence of necessary equipment (SEM). Thus, the present study is limited to providing morphological elements that could help with future descriptions of Clades 4 and 12.

Mega-barcoding provides a powerful method to unravel the underestimated species richness of the *Terebellides* species complex and map its distribution throughout Swedish marine environments. A deeper understanding of *Terebellides* diversity and distribution patterns enables a more robust assessment of their ecological roles within benthic communities and their overall contribution to the health of marine habitats. By integrating DNA barcoding with traditional morphological analysis, the detection of diagnostic traits unique to specific species is greatly facilitated. This synergistic approach significantly enhances the precision and efficiency of environmental monitoring for benthic biodiversity. Among the species identified in this study, two undescribed species underwent a preliminary morphological analysis, which will support their forthcoming formal taxonomic description.

Acknowledgements

I would like to thank my supervisors Maël Grosse and Arne Nygren for their guidance and continuous support throughout this research project. I am also grateful to the staff and researchers at Tjärno Marine Laboratory for providing the facilities required for this study and for fostering a highly stimulating and welcoming environment. I would like to express my sincere thanks to my master's coordinator Catharina Olsson for her valuable assistance and to the Department of Biological and Environmental Sciences for its contribution. I would also like to thank all my

colleagues and friends for enriching discussions and motivation. Special thanks to my family for their outstanding support.

References

Al-Shuhaib, M. B. S., & Hashim, H. O. (2023). Mastering DNA chromatogram analysis in Sanger sequencing for reliable clinical analysis. *Journal of Genetic Engineering and Biotechnology*, *21*(1), 115.

Anderson, M. J., Crist, T. O., Chase, J. M., Vellend, M., Inouye, B. D., Freestone, A. L., ... & Swenson, N. G. (2011). Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecology letters*, *14*(1), 19-28.

Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. *Global ecology and biogeography*, *19*(1), 134-143.

Barroso, M., Moreira, J., Capa, M., Nygren, A., & Parapar, J. (2022). A further step towards the characterisation of Terebellides (Annelida, Trichobranchidae) diversity in the Northeast Atlantic, with the description of a new species. Zookeys, 1132, 85.

Bengtsson, H., & Stevens, R. L. (1996). Heavy-mineral provinces in southern Skagerrak and northern Kattegat.

Bickford, D., Lohman, D. J., Sodhi, N. S., Ng, P. K., Meier, R., Winker, K., ... & Das, I. (2007). Cryptic species as a window on diversity and conservation. *Trends in ecology & evolution*, *22*(3), 148-155.

Brasier, M. J., Wiklund, H., Neal, L., Jeffreys, R., Linse, K., Ruhl, H., & Glover, A. G. (2016). DNA barcoding uncovers cryptic diversity in 50% of deep-sea Antarctic polychaetes. *Royal Society open science*, *3*(11), 160432.

Briers, R. (2006). Ecology: From individuals to ecosystems.

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular biology and evolution*, *37*(5), 1530-1534.

Burrows, M. T., Schoeman, D. S., Buckley, L. B., Moore, P., Poloczanska, E. S., Brander, K. M., ... & Richardson, A. J. (2011). The pace of shifting climate in marine and terrestrial ecosystems. *Science*, *334*(6056), 652-655.

Costanza, R., d'Arge, R., De Groot, R., Farber, S., Grasso, M., Hannon, B., ... & Van Den Belt, M. (1997). The value of the world's ecosystem services and natural capital. *nature*, *387*(6630), 253-260.

Chang, J. J. M., Ip, Y. C. A., Ng, C. S. L., & Huang, D. (2020). Takeaways from mobile DNA barcoding with BentoLab and MinION. *Genes*, *11*(10), 1121.

Chua, P. Y., Bourlat, S. J., Ferguson, C., Korlevic, P., Zhao, L., Ekrem, T., ... & Lawniczak, M. K. (2023). Future of DNA-based insect monitoring. *Trends in Genetics*, *39*(7), 531-544.

Crain, C. M., Kroeker, K., & Halpern, B. S. (2008). Interactive and cumulative effects of multiple human stressors in marine systems. *Ecology letters*, *11*(12), 1304-1315.

Crossley, B. M., Bai, J., Glaser, A., Maes, R., Porter, E., Killian, M. L., ... & Toohey-Kurth, K. (2020). Guidelines for Sanger sequencing and molecular assay monitoring. *Journal of Veterinary Diagnostic Investigation*, *32*(6), 767-775.

Danovaro, R., Gambi, C., Dell'Anno, A., Corinaldesi, C., Fraschetti, S., Vanreusel, A., ... & Gooday, A. J. (2008). Exponential decline of deep-sea ecosystem functioning linked to benthic biodiversity loss. *Current Biology*, *18*(1), 1-8.

De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic biology*, *56*(6), 879-886.

Doney, S. C., Ruckelshaus, M., Duffy, J. E., Barry, J. P., Chan, F., English, C. A., ... & Talley, L. D. (2012). Climate change impacts on marine ecosystems. *Annual review of marine science*, *4*(2012), 11-37.

Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: improving the ultrafast bootstrap approximation. *Molecular biology and evolution*, *35*(2), 518-522.

Draper, N. (2008). Identification of SNPs, or mutations in sequence chromatograms. In *Genomics Protocols* (pp. 35-52). Totowa, NJ: Humana Press.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, *32*(5), 1792-1797.

English, O. (1992). Convention on biological diversity. *Estud Av*, *20*(364), 10-1017.

Ewing, B., Hillier, L., Wendl, M. C., & Green, P. (1998). Base-calling of automated sequencer traces usingPhred. I. Accuracy assessment. *Genome research*, *8*(3), 175-185.

Ewing, B., & Green, P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome research*, *8*(3), 186-194.

Eurofins Genomics. (2025). *Sanger sequencing services and PlateSeq kits pricing*. https://eurofinsgenomics.eu/en/custom-dna-sequencing/eurofins-services/plateseq-kits/

Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular marine biology and biotechnology*, *3*(5), 294–299.

The Galaxy Community. (2024). The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. *Nucleic Acids Research*, *52*(W1), W83–W94.

Halpern, B. S., Walbridge, S., Selkoe, K. A., Kappel, C. V., Micheli, F., d'Agrosa, C., ... & Watson, R. (2008). A global map of human impact on marine ecosystems. *science*, *319*(5865), 948-952.

Hauer, C., Leitner, P., Unfer, G., Pulg, U., Habersack, H., & Graf, W. (2018). The role of sediment and sediment dynamics in the aquatic environment. In *Riverine ecosystem management: Science for governing towards a sustainable future* (pp. 151-169). Cham: Springer International Publishing.

Hebert, P. D., Cywinska, A., Ball, S. L., & DeWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1512), 313-321.

Helgason, G. V. (2011). Taxonomy and distribution of Terebellides (Polychaeta, Trichobranchidae) in Icelandic waters, with the description of a new species.

Hoegh-Guldberg, O., & Bruno, J. F. (2010). The impact of climate change on the world's marine ecosystems. *Science*, *328*(5985), 1523-1528.

Hutchings, P. (1998). Biodiversity and functioning of polychaetes in benthic sediments. *Biodiversity & Conservation*, *7*, 1133-1145.

Hutchings, P. A., Nogueira, J. M. M., Carrerette, O., Purschke, G., & Westheide, W. (2017). Terebelliformia. *The Handbook of Zoology Online*.

Jost, L. (2007). Partitioning diversity into independent alpha and beta components. *Ecology*, *88*(10), 2427-2439.

Jurasinski, G., Retzer, V., & Beierkuhnlein, C. (2009). Inventory, differentiation, and proportional diversity: a consistent terminology for quantifying species diversity. *Oecologia*, *159*(1), 15-26.

Koleff, P., Gaston, K. J., & Lennon, J. J. (2003). Measuring beta diversity for presence–absence data. *Journal of Animal Ecology*, *72*(3), 367-382.

Kress, W. J., & Erickson, D. L. (2012). *DNA barcodes: methods and protocols* (pp. 3-8). Humana Press.

Kritzer, J. P., DeLucia, M. B., Greene, E., Shumway, C., Topolski, M. F., Thomas-Blate, J., ... & Smith, K. (2016). The importance of benthic habitats for coastal fisheries. *BioScience*, *66*(4), 274-284.

Kudryavtseva, A. A., Novoyatlova, U. S., Chuyko, A., Gaeva, D. R., Vlasov, A. V., & Manukhov, I. V. (2023). Disjunct habitat of cryptic Terebellides (Annelida, Trichobranchidae) species shows a phylogenetic link between polychaetes from the White and the North Seas. *Scientific Reports*, *13*(1), 22926.

Larsson, A. (2014). AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics*, *30*(22), 3276-3278.

Lavesque, N., Hutchings, P., Daffe, G., Nygren, A., & Londoño-Mesa, M. H. (2019). A revision of the French Trichobranchidae (Polychaeta), with descriptions of nine new species. *Zootaxa*, *4664*(2), zootaxa-4664.

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., ... & Machida, R. J. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Frontiers in zoology*, *10*(1), 34.

Letunic, I., & Bork, P. (2024). Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic acids research*, *52*(W1), W78-W82.

Lovett, G. M., Burns, D. A., Driscoll, C. T., Jenkins, J. C., Mitchell, M. J., Rustad, L., ... & Haeuber, R. (2007). Who needs environmental monitoring?. *Frontiers in Ecology and the Environment*, *5*(5), 253-260.

Magurran, A. E. (2003). *Measuring biological diversity*. John Wiley & Sons.

Microsoft Corporation. (2018). *Microsoft Excel* (Version 16.16.2) [Computer software]. https://www.microsoft.com/en-us/microsoft-365/excel

Miller, S. A., & Harley, J. P. (2016). *Zoology* (10th ed.). McGraw-Hill Education.

Noss, R. F. (1990). Indicators for monitoring biodiversity: a hierarchical approach. *Conservation biology*, *4*(4), 355-364.

Nygren, A. (2014). Cryptic polychaete diversity: a review. *Zoologica Scripta*, *43*(2), 172-183.

Nygren, A., Parapar, J., Pons, J., Meißner, K., Bakken, T., Kongsrud, J. A., ... & Capa, M. (2018). A mega-cryptic species complex hidden among one of the most common annelids in the North East Atlantic. *PLoS One*, *13*(6), e0198356.

Palumbi, S. R. (1994). Genetic divergence, reproductive isolation, and marine speciation. *Annual review of ecology and systematics*, 547-572.

Palumbi, S. R. (2001). Humans as the world's greatest evolutionary force. *Science*, *293*(5536), 1786-1790.

Palumbi, S. R., McLeod, K. L., & Grünbaum, D. (2008). Ecosystems in action: lessons from marine ecology about recovery, resistance, and reversibility. *BioScience*, *58*(1), 33-42.

Palumbi, S. R., Sandifer, P. A., Allan, J. D., Beck, M. W., Fautin, D. G., Fogarty, M. J., ... & Wall, D. H. (2009). Managing for ocean biodiversity to sustain marine ecosystem services. *Frontiers in Ecology and the Environment*, *7*(4), 204-211.

Parapar, J., Moreira, J., & Helgason, G. V. (2011). Taxonomy and distribution of Terebellides (Polychaeta, Trichobranchidae) in Icelandic waters, with the description of a new species. *Zootaxa*, *2983*(1), 1-20.

Parapar, J., Moreira, J., & O'Reilly, M. (2016c). A new species of Terebellides (Polychaeta: Trichobranchidae) from Scottish waters with an insight into branchial morphology. Marine Biodiversity, 46(1), 211-225.

Parapar, J., Capa, M., Nygren, A., & Moreira, J. (2020a). To name but a few: descriptions of five new species of Terebellides (Annelida, Trichobranchidae) from the North East Atlantic. ZooKeys, 992, 1.

Parapar, J., Martin, D., & Moreira, J. (2020b). On the diversity of Terebellides (Annelida, Trichobranchidae) in West Africa, seven new species and the redescription of T. africana Augener, 1918 stat. prom. Zootaxa, 4771(1), zootaxa-4771.

Peterson, C. H., & Lubchenco, J. (1997). Marine ecosystem services.

Pomerantz, A., Peñafiel, N., Arteaga, A., Bustamante, L., Pichardo, F., Coloma, L. A., ... & Prost, S. (2018). Real-time DNA barcoding in a rainforest using nanopore sequencing: opportunities for rapid biodiversity assessments and local capacity building. *GigaScience*, *7*(4), giy033.

Pomerantz, A., Sahlin, K., Vasiljevic, N., Seah, A., Lim, M., Humble, E., ... & Prost, S. (2022). Rapid in situ identification of biological specimens via DNA amplicon sequencing using miniaturized laboratory equipment. *Nature Protocols*, *17*(6), 1415-1443.

QGIS.org. (2026). *QGIS Geographic Information System* (Version 3.40.5-Bratislava) [Computer software]. QGIS Association. http://www.qgis.org

Ratnasingham, S., & Hebert, P. D. (2007). BOLD: The Barcode of Life Data System (http://www. barcodinglife. org). *Molecular ecology notes*, *7*(3), 355-364.

Rouse, G., Pleijel, F., & Tilic, E. (2022). *Annelida*. Oxford University Press.

Saccone, C., De Giorgi, C., Gissi, C., Pesole, G., & Reyes, A. (1999). Evolutionary genomics in Metazoa: the mitochondrial DNA as a model system. *Gene*, *238*(1), 195-209.

Savolainen, V., Cowan, R. S., Vogler, A. P., Roderick, G. K., & Lane, R. (2005). Towards writing the encyclopaedia of life: an introduction to DNA barcoding.*Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1462), 1805-1811.

Schueller, M., & Hutchings, P. A. (2010). New insights in the taxonomy of Trichobranchidae (Polychaeta) with description of a new Terebellides species from Australia. Zootaxa, 2395(1), 1-16.

Schueller, M., & Hutchings, P. A. (2013). New species of Terebellides (Polychaeta: Trichobranchidae) from the deep Southern Ocean, with a key to all described species. *Zootaxa*, *3619*, 1-45.

Selig, E. R., Hole, D. G., Allison, E. H., Arkema, K. K., McKinnon, M. C., Chu, J., ... & Zvoleff, A. (2019). Mapping global human dependence on marine ecosystems. *Conservation Letters*, *12*(2), e12617.

Shendure, J., & Ji, H. (2008). Next-generation DNA sequencing. *Nature biotechnology*, *26*(10), 1135-1145.

SLU Artdatabanken (2025). *Artfakta: Terebellides*. https://artfakta.se/taxa/1007592 [2025-02-02]

Smith, T. M., & Smith, R. L. (2012). Elements of Ecology, eight editions. *Person Benjamin Cummings: San Francisco*, 1-611.

Socolar, J. B., Gilroy, J. J., Kunin, W. E., & Edwards, D. P. (2016). How should beta-diversity inform biodiversity conservation?. *Trends in ecology & evolution*, *31*(1), 67-80.

Srivathsan, A., Hartop, E., Puniamoorthy, J., Lee, W. T., Kutty, S. N., Kurina, O., & Meier, R. (2019). Rapid, large-scale species discovery in hyperdiverse taxa using 1D MinION sequencing. *BMC biology*, *17*(1), 96.

Srivathsan, A., Lee, L., Katoh, K., Hartop, E., Kutty, S. N., Wong, J., ... & Meier, R. (2021). ONTbarcoder and MinION barcodes aid biodiversity discovery and identification by everyone, for everyone. *BMC biology*, *19*(1), 217.

Srivathsan, A., Feng, V., Suárez, D., Emerson, B., & Meier, R. (2024). ONTbarcoder 2.0: rapid species discovery and identification with real-time barcoding facilitated by Oxford Nanopore R10. 4. *Cladistics*, *40*(2), 192-203.

Srivathsan, A., & Meier, R. (2024). Scalable, cost-effective, and decentralized DNA barcoding with Oxford nanopore sequencing. In *DNA barcoding: Methods and protocols* (pp. 223-238). New York, NY: Springer US.

Stanhope, M. J., Connelly, M. M., & Hartwick, B. (1992). Evolution of a crustacean chemical communication channel: behavioral and ecological genetic evidence for a habitat-modified, race-specific pheromone. *Journal of chemical ecology*, *18*, 1871-1887.

Struck, T. H., Paul, C., Hill, N., Hartmann, S., Hösel, C., Kube, M., ... & Bleidorn, C. (2011). Phylogenomic analyses unravel annelid evolution. *Nature*, *471*(7336), 95-98.

Struck, T. H., Purschke, G., Dordel, J., Hösel, C., Nesnidal, M. P., Diersing, F., ... & Hartmann, S. (2014). Phylogeny and evolution of Annelida based on molecular data. *Deep Metazoan Phylogeny: The Backbone of the Tree of Life. New Insights from Analyses of Molecules, Morphology, and Theory of Data Analysis*, 143-160.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., Von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature methods*, *14*(6), 587-589.

Svansson, A. (1975). Physical and chemical oceanography of the Skagerrak and the Kattegat.

Tuomisto, H. (2010a). A diversity of beta diversities: straightening up a concept gone awry. Part 1. Defining beta diversity as a function of alpha and gamma diversity. *Ecography*, *33*(1), 2-22.

Tuomisto, H. (2010b). A diversity of beta diversities: straightening up a concept gone awry. Part 2. Quantifying beta diversity and related phenomena. *Ecography*, *33*(1), 23-45.

Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., & Orr, R. B. (2020). *Campbell biology* (12th ed.). Pearson.

Vellend, M. (2010). Conceptual synthesis in community ecology. *The Quarterly review of biology*, *85*(2), 183-206.

Vellend, M. (2001). Do commonly used indices of β-diversity measure species turnover?. *Journal of Vegetation Science*, *12*(4), 545-552.

Watsa, M., Erkenswick, G. A., Pomerantz, A., & Prost, S. (2020). Portable sequencing as a teaching tool in conservation and biodiversity research. *PLoS biology*, *18*(4), e3000667.

Whittaker, R. H. (1960). Vegetation of the Siskiyou mountains, Oregon and California. *Ecological monographs*, *30*(3), 279-338.

Whittaker, R. H. (1972). Evolution and measurement of species diversity. *Taxon*, *21*(2-3), 213-251.

Wilson, E. O. (1985). The biological diversity crisis. *BioScience*, *35*(11), 700-706.

Wilson, E. O. (1988). The current state of biological diversity. *Biodiversity*, *521*(1), 3-18.

Winston, J. E., & Disney, H. (2000). Describing species: practical taxonomic procedure for biologists. *Nature*, *405*(6787), 619.

Worm, B., Barbier, E. B., Beaumont, N., Duffy, J. E., Folke, C., Halpern, B. S., ... & Watson, R. (2006). Impacts of biodiversity loss on ocean ecosystem services. *science*, *314*(5800), 787-790.

Appendix 1

POPULAR SCIENCE SUMMARY

To date, while roughly 2 million species (ranging from 1.2 to 1.8 million) have been inventoried, current estimations suggest a global species diversity ranging from 8 to 20 million on Earth! Besides large mammals that have largely been documented (and which no longer really have a place to hide!), the vast majority of life on Earth is still a mystery waiting to be uncovered. This raises crucial questions: what factors hinder species recognition, and why is their identification so important?

Taxonomy & species concept

Before we start, it is necessary to address the scientific field that deals with the identification of species: taxonomy, which is the classification of formally named groups of organisms. Taxonomists perform the identification, the classification and the naming of organisms by using both morphological and molecular (DNA sequences) data. However, one first obstacle encountered by taxonomists is assenting on what definition of 'species' they base themselves on. What is a species exactly? There is no single unanimous agreement and authors are still arguing over which exact definition to adopt. The most recurrent definition used is the 'Biological species concept' (BSC). This concept defines a species as a population (or group of populations) of actually or potentially interbreeding organisms that are reproductively isolated from other such groups.

Many overlooked species : why?

Why is it that it can be difficult to easily identify species? There are several reasons for this. Besides the fact that some natural environments are difficult to access, sometimes inaccessible, most of species are predominantly tiny organisms, the vast majority of which remain invisible to the naked eye. Moreover, species are sometimes overlooked because of outdated taxonomic information or incomplete distribution data. Finally, the allocation of conservation resources and research funding is often skewed towards aesthetically appealing taxa, leaving less charismatic groups understudied and unprotected. Finally, another specific and compelling reason behind the challenges of species identification will hold our attention for the remainder of our narrative...

Cryptic species

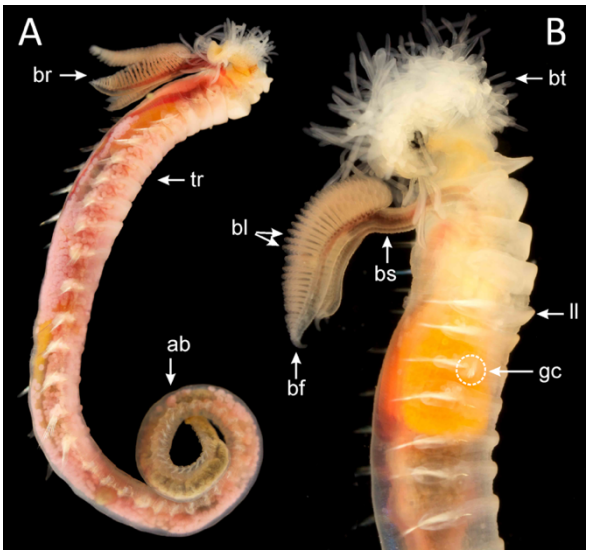
One of the major puzzles encountered among taxonomists, even before the Linnaean nomenclature was adopted, is the identification of cryptic species. Cryptic species are defined as 'two or more distinct species that are erroneously classified (and thus hidden) under one species name'. What makes them cryptic? Among which groups and in which natural environments are they mostly hidden? First of all, it is necessary to bear in mind that taxonomy has long been based solely on morphological descriptions, biased by our human perception of the living world, which is essentially depicted by our visual sense. However, not all species rely only on visual perception. Indeed, one reason that would lead to the emergence of cryptic species is the use of non-visual mating signals among certain taxa, such as acoustic (sound, vibration), chemical (pheromones) or even electrical signals, making highly differentiated morphology unnecessary. Another explanation concerns taxa undergoing particularly strong selection pressure due to adaptation to a specific host (parasites) or to extreme environments (extremophiles). Indeed, extreme conditions reduce the range of adaptive possibilities and can lead to evolutionary convergences in terms of morphology (evolutionary

convergence is the independent evolution of similar traits in distantly related organisms). Then, even though some studies conclude the opposite, certain environments harbor more cryptic species than others. Indeed, we would expect that they would flourish in ecosystems where visual communication signals are obstructed and would force them to use other means of communication, as seen previously. Thus, two types of environment quickly come to mind, the tropical rainforest alongside another environment that will hold our attention from now on…

Marine habitats & bristle worms

Oceans remain the most unexplored environment on Earth, partly because of their vastness, but also due to the significant challenges humans face in exploring their depths.

Annelids are segmented worms found globally in aquatic and terrestrial environments and are abundant in diverse marine habitats (hydrothermal vents to intertidal zones). Traditionally divided into Polychaeta (marine worms) and Clitellata (earthworms, leeches), phylogenomic analyses now recognize Errantia (mainly 'mobile' annelids) and Sedentaria (mainly annelids fixed to the substrate) as major clades. Marine annelids play a substantial role in benthic communities, as they are either predators, herbivores, scavengers or filter feeders. The genus *Terebellides* (Annelida) comprises tube-dwelling deposit feeders mainly on continental shelves and slopes and is the most diverse genus within Trichobranchidae (52 species). In the North East Atlantic, morphology-based studies reported seven species, including two on Swedish coasts (*Terebellides stroemii* & *Terebellides gracilis*). However, morphology alone often underestimates marine annelid diversity due to cryptic species.



Specimen of *Terebellides* © 2018 Nygren *et al.*

Ongoing study

Genetic analyses of *Terebellides* in the North East Atlantic revealed 25 species instead of seven, with ten in the Skagerrak and Kattegat (only four described). Therefore, genetic data is essential for accurate *Terebellides* identification in Sweden, a frequently encountered group crucial for environmental monitoring.

Why is species identification important?

Incorrect identification of species complexes (group of many cryptic species) negatively impacts species conservation, habitat protection, and diverse fields including pest/pathogen control, human health (parasites, vectors), and bioprospecting. In conservation, accurate identification is crucial for effective habitat management via precise distribution mapping of endangered species. This knowledge also improves ecotoxicological assessments and the utility of bioindicator species in environmental monitoring ((bioindicator species provide information about an ecosystem's health). As *Terebellides* is frequently found off Sweden's coasts, its accurate identification is essential for these monitoring efforts.