Mitochondrial cytochrome c oxidase subunit I (COI) metabarcoding of Foraminifera communities using taxon-specific primers

Jan-Niklas Macher[1], Dimitra Maria Bloska[1], Maria Holzmann[2], Elsa B. Girard[1,3], Jan Pawlowski[4] and Willem Renema[1,3]

[1] Marine Biodiversity, Naturalis Biodiversity Center, Leiden, The Netherlands
[2] Department of Genetics & Evolution, University of Geneva, Geneva, Switzerland
[3] Department of Ecosystem & Landscape Dynamics, University of Amsterdam, Amsterdam, Netherlands
[4] Laboratory of Paleoceanography, Institute of Oceanology Polish Academy of Sciences, Sopot, Poland

ABSTRACT

Foraminifera are a species-rich phylum of rhizarian protists that are highly abundant in most marine environments. Molecular methods such as metabarcoding have revealed a high, yet undescribed diversity of Foraminifera. However, so far only one molecular marker, the 18S ribosomal RNA, was available for metabarcoding studies on Foraminifera. Primers that allow amplification of foraminiferal mitochondrial cytochrome oxidase I (COI) and identification of Foraminifera species were recently published. Here we test the performance of these primers for the amplification of whole foraminiferal communities, and compare their performance to that of the highly degenerate LerayXT primers, which amplify the same COI region in a wide range of eukaryotes. We applied metabarcoding to 48 samples taken along three transects spanning a North Sea beach in the Netherlands from dunes to the low tide level, and analysed both sediment samples and meiofauna samples, which contained taxa between 42 μm and 1 mm in body size obtained by decantation from sand samples. We used single-cell metabarcoding (*Girard et al., 2022*) to generate a COI reference library containing 32 species of Foraminifera, and used this to taxonomically annotate our community metabarcoding data. Our analyses show that the highly degenerate LerayXT primers do not amplify Foraminifera, while the Foraminifera primers are highly Foraminifera- specific, with about 90% of reads assigned to Foraminifera and amplifying taxa from all major groups, *i.e.*, monothalamids, Globothalamea, and Tubothalamea. We identified 176 Foraminifera ASVs and found a change in Foraminifera community composition along the beach transects from high tide to low tide level, and a dominance of single-chambered monothalamid Foraminifera. Our results highlight that COI metabarcoding can be a powerful tool for assessing Foraminiferal communities.

Subjects Biodiversity, Ecology, Marine Biology, Molecular Biology
Keywords Foraminifera, Metabarcoding, Beach, Community composition, Intertidal, Molecular biodiversity

INTRODUCTION

Foraminifera are a species-rich phylum of rhizarian protists (*Goldstein, 1999*; *Gupta, 2003*; *Burki et al., 2010*) found mostly in marine environments. Studies on species diversity and ecology of Foraminifera are conducted to understand both past (*Keller, 1983*; *Charnock & Jones, 1990*; *Scheibner, Speijer & Marzouk, 2005*) and recent (*Hallock et al., 2003*; *Murray, 2006*; *Nooijer et al., 2008*; *Pawlowski et al., 2014*) ecosystems. Foraminifera can be identified on both morphological characters and molecular data, which allows deeper insight into their diversity (*Pawlowski & Holzmann, 2014*; *Pawlowski, Lejzerowicz & Esling, 2014*; *Morard et al., 2015*; *Morard et al., 2016*). Molecular work on foraminiferal species has provided insights into their phylogeny (*Holzmann et al., 2001*; *Pawlowski & Holzmann, 2002*; *Darling & Wade, 2008*; *Pawlowski, Holzmann & Tyszka, 2013*; *Holzmann & Pawlowski, 2017*) and led to the discovery of cryptic diversity in widely distributed morphospecies (*Darling, Kucera & Wade, 2007*; *Morard et al., 2016*; *Prazeres et al., 2020*; *Macher et al., 2021a*). Molecular metabarcoding, *i.e.*, the amplification, sequencing and analysis of whole communities (*Taberlet et al., 2012*), is commonly applied to Foraminifera communities to detect species diversity in a wide range of ecosystems (*Pawlowski, Lejzerowicz & Esling, 2014*; *Morard et al., 2019*; *Holzmann et al., 2021*), for environmental impact assessments (*Pawlowski et al., 2014*; *Laroche et al., 2016*; *Frontalini et al., 2018*) and to study the ecology of Foraminifera (*He et al., 2019*; *Chronopoulou et al., 2019*; *Greco, Morard & Kucera, 2020*). Until recently, only one genetic marker, the nuclear 18S ribosomal RNA (the small subunit ribosomal ribonucleic acid; SSU), has been used for barcoding and metabarcoding of Foraminifera. The 18S rRNA allows identification of most foraminiferal species (*Pawlowski & Holzmann, 2014*; *Morard et al., 2019*), but some species show minimal or no variability (*Schweizer et al., 2011*; *Borrelli et al., 2018*) and others show hypervariability (*Weber & Pawlowski, 2014*; *Morard et al., 2016*; *Prazeres et al., 2020*), which can hamper species identification. To further advance molecular studies on Foraminifera, additional and easily obtainable molecular markers are desirable. This could potentially improve species identification, phylogenetic analyses and ecological studies. Recently, the first mitochondrial cytochrome c oxidase subunit I (COI) gene sequences of Foraminifera were published, together with primers which amplify a COI fragment (*Macher et al., 2021b*; *Girard et al., 2022*) overlapping the established Leray fragment, which is used in metabarcoding studies on a wide range of Eukaryota (*Leray et al., 2013*; *Wangensteen et al., 2018*). COI is a promising molecular marker for species identification in Foraminifera, but so far the newly developed COI primers were only applied to single foraminiferal specimens for the purpose of reference barcoding and assessment of marker variability (*Macher et al., 2021b*; *Girard et al., 2022*). Here we test whether the new Foraminifera COI primers can be used for metabarcoding of Foraminifera communities from environmental samples.

We applied metabarcoding to 48 samples taken along three intertidal transects spanning a North Sea beach in the Netherlands from dunes to the low tide line. We test taxon specificity of the Foraminifera primers, *i.e.*, whether they amplify mostly Foraminifera or also a wide range of other taxa as reported for several degenerate COI metabarcoding

primers (*Weigand & Macher, 2018*; *Wangensteen et al., 2018*). We compare the performance of the Foraminifera COI primers to that of the highly degenerate LerayXT primers, which are known to amplify COI of a wide range of eukaryotes and even prokaryotes (*Collins et al., 2019*). Second, we test whether the inferred community composition of Foraminifera is influenced by the sampling method (sediment core vs. meiofauna sample), and third, whether the metabarcoding data can be used to infer changes in foraminiferal communities along the beach transects from high tide to low tide line.

## MATERIALS & METHODS

Field work

We took samples along three parallel transects (3 m apart) from the dunes to the low-tide line at Katwijk in the Netherlands (coordinates: 52.188479 N, 4.376894 E) on 13th August 2021 at low tide. We sampled eight sites along each of the transects: the first sample was taken at the foot of the dunes, above the high-tide line, the second sample halfway between the dunes and the high-tide line, and the remaining six samples equidistantly spaced from high-tide line to low-tide line (see Fig. 1). Since we took samples from three parallel transects, we sampled a total of 24 sites. Samples were taken and the beach state was recorded according to protocols for sampling the intertidal zone of sandy beaches, including measurement of the beach slope and assessing the breaker height (*McLachlan, Defeo & Short, 2018*).

We collected two types of samples per sampling site: one sediment core of 5 cm diameter and a length of 10 cm (volume ≈200 ml), and a second sediment core of 1 cm diameter and a length of 10 cm (volume ≈8 ml), with standard, sterile plastic syringes cut open at the front end. Immediately after sampling, we transferred the small (≈8 ml) sediment core to a 50 ml falcon tube and preserved the sample with 30 ml of 96% EtOH. We extracted meiofauna from the larger sediment core (≈200 ml) by adding 500 ml of $MgCl_2$ solution to the sediment in a sterile bottle, which anaesthetises meiofauna and allows subsequent separation from the sediment by decantation (*Somerfield, Warwick & Moens, 2005*). After 5 min, samples in $MgCL_2$ solution were carefully swirled and the supernatant containing the meiofauna decanted through a 1 mm and 41 um sieve cascade, as commonly done in beach meiofauna studies (*Haenel et al., 2017*; *Martínez et al., 2020*; *Castro et al., 2021*; *Gielings et al., 2021*). The meiofauna fraction retained on the 41um sieve was rinsed into sterile 15 ml Falcon tubes and preserved with 10 ml 96% EtOH. All samples were transported back to the Naturalis Biodiversity Centre laboratory and stored at −20 °C until further processing.

DNA extraction

Ethanol from Falcon tubes containing the meiofauna samples was evaporated at 50 °C overnight in a sterile warming cabinet, and we transferred the dried samples to 2 ml Eppendorf tubes. We extracted DNA with the Macherey Nagel NucleoSpin Soil kit (Macherey Nagel, Düren, Germany) according to the standard protocol including bead beating, but starting with an additional overnight Proteinase K digestion step (50 ul 250 µg/ml protK (Thermo Fisher Scientific, Waltham, USA) added to the lysis buffer provided with the kit) to improve cell lysis, as done in previous studies on sediment meiofauna

Katwijk Beach
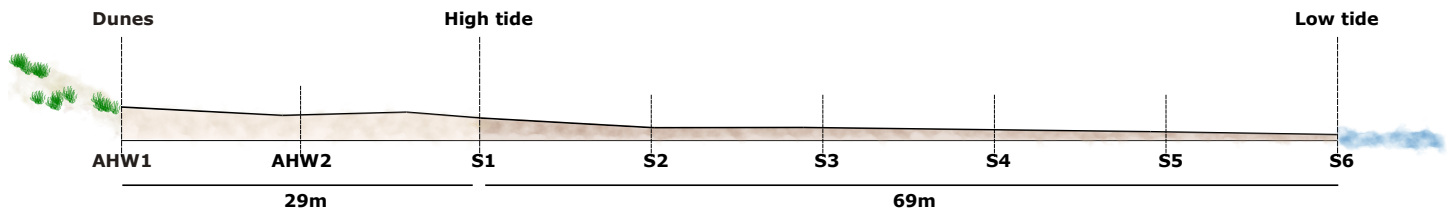
Figure 1 **Profile of Katwijk beach, showing the eight sampling points along the transect from dunes to the low tide line.** AHW1-AHW2: Sampling sites above the high water line; S1–S6: Sampling sites in the intertidal area.

Full-size 🖼 DOI: 10.7717/peerj.13952/fig-1

(*Weigand & Macher, 2018*; *Martínez et al., 2020*). The sediment samples were transferred to sterile petri dishes and dried overnight at 50 °C in a sterile warming cabinet. After drying, we homogenised each sample with a sterile spatula, and a subsample of 0.5 g was transferred to a 2 ml Eppendorf tube. The extraction followed the protocol as described above for meiofauna samples.

Amplification of COI for community metabarcoding

We amplified both sediment and meiofauna samples with a two-step PCR protocol, with the commonly used LerayXT primers targeting a wide range of Eukaryota (*Wangensteen et al., 2018*; *Collins et al., 2019*), and the recently published Foraminifera COI primers (Foraminifera_COI_fwd1: 5′-GWGGWGTTAATGCTGGTYGAAC -3′, Foraminifera_COI_rev: 5′- RWRCTTCWGGATGWCTAAGARATC-3′) targeting the same COI region (*Macher et al., 2021b*) with an amplicon length of 310 to 320 base pairs. For the first PCR, each reaction contained 11.7ul mQ water, 2ul Qiagen CL buffer (10x; Qiagen, Hilden, Germany), 0.4 ul MgCL2 (25 mM; Qiagen), 0.8ul Bovine Serum Albumine (BSA, 10 mg/ml), 0.4 ul dNTPs (2.5 mM), 0.2ul Qiagen Taq (5U/ul), 1ul of each nextera-tailed primer (10 pMol/ul), and 2.5 ul of DNA template. PCR amplification was performed with 3 min of initial denaturation at 96 °C, 30 cycles of denaturation for 15 s at 96 °C, annealing at 50 °C for 30 s, extension for 40 s at 72 °C, followed by a final extension at 72 °C for 5 min. Three negative controls (Milli-Q water, Merck, Kenilworth, USA) were processed together with the samples to check for potential contamination. After the first PCR, samples were cleaned with AMPure beads (Beckman Coulter, Brea, United States) at a ratio of 0.9:1 according to protocol. For the second PCR, samples were amplified with individually tagged primers following the same protocol as above, using the PCR product from the first PCR as template and with PCR cycle number reduced to 10. We measured DNA concentrations using the TapeStation (Agilent Technologies. Santa Clara, CA, USA) with the High Sensitivity Kit and equimolarly pooled samples. The final library was cleaned with AMPure beads as described above and sent for sequencing on part of an Illumina MiSeq run (2 × 300 bp read length) at Baseclear (Leiden, The Netherlands).

Sequencing of Foraminifera reference species

We performed single-cell metabarcoding on 36 specimens from 32 morphospecies of Foraminifera as a reference for taxonomic annotation, since only a limited number of foraminiferal COI references are available in the NCBI reference databases to date (*Macher et al., 2021b*). The DNA was obtained from the existing Foraminifera DNA collection of the Department of Genetics and Evolution, University of Geneva, curated by the co-authors Maria Holzmann and Jan Pawlowski. We amplified and sequenced the specimens using the Foram_COI_fwd1/ Foram_COI_rev primers with the metabarcoding protocol described above, with the following differences in PCR protocol: The DNA template for the first PCR was 10x diluted, and we used 40 cycles of amplification for the first PCR and 8 cycles for the tagging PCR. We measured DNA concentrations using the TapeStation (Agilent Technologies, Santa Clara, USA) and equimolarly pooled samples. Sequencing was conducted on part of a Illumina MiSeq run (2×300 bp) at Baseclear, Leiden. Bioinformatic processing was performed with the settings described below for community metabarcoding data. We performed a contamination check using the 'decontam' pipeline (v.1.14.0) (*Davis et al., 2018*) in R (v.4.1.2) using the package 'phyloseq' (v.1.38.0) (*McMurdie & Holmes, 2013*). To remove any potential low-key contaminations not picked up by the 'decontam' pipeline, we further retained only those ASVs with a read abundance >50% per specimen to obtain reliable references. We added 44 COI sequences of 16 previously sequenced species (*Macher et al., 2021b*) to the reference database to increase taxonomic coverage. To further increase taxonomic coverage, we amplified and Sanger sequenced 21 specimens from six morphospecies, with the same primers and the protocol described for COI barcoding of Foraminifera (*Macher et al., 2021b*). We quality trimmed and aligned these sequences in MEGA 11 (*Tamura, Stecher & Kumar, 2021*). One specimen (*Glabratella* sp. 17919) was excluded from further analyses due to low read quality and the sequence being shorter than expected.

We aligned all reference sequences using MAFFT (v. 1.4.0) (*Katoh et al., 2002*) as implemented in Geneious (v2020.2) and translated the sequences using mitochondrial translation table 4 to check for potential stop codons that might indicate sequencing errors or NUMTs (*Macher et al., 2021b*; *Girard et al., 2022*). This led to the exclusion of the sequences of *Psammophaga* sp. 19297, *Psammophaga* sp. 19263 and *Psammophaga* sp. 19299, which showed internal stop codons. As an additional quality check and to assess whether the newly generated reference sequences cluster in the expected phylogenetic position of a species based on previous studies, we calculated a cladogram using FastTree (*Price, Dehal & Arkin, 2010*) as implemented in Geneious Prime (v. 2021.2). Taxonomic assignment of the reference sequences followed (*Holzmann et al., 2022*; *Holzmann et al., 2021*; *Pawlowski, Holzmann & Tyszka, 2013*; *Holzmann & Pawlowski, 2017*; *Siemensma et al., 2017*; *Renema, 2018*). Furthermore, as in previous studies (*Macher et al., 2021b*; *Girard et al., 2022*), we used the ASAP species delimitation algorithm (*Puillandre, Brouillet & Achaz, 2021*) with default Kimura K80 settings on the newly created reference database to test which genetic distance is appropriate for defining species based on the amplified COI fragment.

Bioinformatic processing of community metabarcoding data

We processed metabarcoding raw reads using the Galaxy platform (*Afgan et al., 2018*) following a standard workflow (*Beentjes et al., 2019*; *Macher et al., 2020*; *Girard et al., 2022*). The LerayXT and the Foraminifera dataset were processed separately. FLASH (v1.2.11) (*Magoč & Salzberg, 2011*) was used to merge reads with a minimum overlap of 150, a maximum overlap of 300, and a maximum mismatch ratio of 0.2 allowed. Non-merged reads were discarded. Cutadapt (v.2.8) (*Martin, 2011*) was used to trim primers (settings: both primers need to be present, minimum number of matching bases 10, maximum error rate 0.2). PrinSeq (v.0.20.4) (*Schmieder & Edwards, 2011*) was used to filter and trim sequences to 300 base pairs to remove reads that contain gaps or indels, which can be present due to sequencing errors, NUMTs or amplification of non-eukaryotic taxa (*Wangensteen et al., 2018*; *Macher et al., 2018*; *Collins et al., 2019*). UNOISE implemented in USEARCH (v10.0.240) (*Edgar, 2016*) was used for clustering of sequences into Amplicon Sequence Variants (ASVs) with thresholds of alpha = 4 and a minimum number of eight reads for the denoising approach, which is similar to previously tested settings (*Turon et al., 2019*). In order to remove potential spurious sequences present due to low-key contamination or tag switching (*Schnell, Bohmann & Gilbert, 2015*), we removed samples with less than 10,000 reads from the dataset, and retained only ASVs with more than 0.01% read abundance per sample.

We annotated ASVs to taxonomic names using the BOLDigger tool (v.1.2.6) (*Buchner & Leese, 2020*), which accesses both public and private references deposited in the Barcode of Life Database (*Ratnasingham & Hebert, 2007*). For taxonomic identification of Foraminifera, both the LerayXT and Foraminifera dataset were additionally searched against the newly constructed reference database of Foraminifera COI sequences described above using blastn (*Altschul et al., 1990*), with the algorithm set to retrieve annotations for sequences with a minimum of 80% sequence identity and 90% query coverage to a reference. The annotations were added to the taxonomy table obtained from BOLDigger. Based on the ASAP species delimitation results we applied to the newly created reference database, and previous results (*Girard et al., 2022*), we used a threshold of 99% identity for species level matches. For higher taxonomic ranks we used identity thresholds that are similar or slightly stricter than those previously used used for assigning taxonomic ranks in unknown eukaryotic communities (*Holovachov, 2016*; *Elbrecht et al., 2017*; *Weigand & Macher, 2018*): genus 96%; family 94%; order 90%; class 85%; phylum 80%.

Taxon specificity of Foram primers and ASV richness

We assessed the taxon-specificity of the COI primers by calculating the fraction of reads annotated to Foraminifera in the two datasets, *i.e.*, samples amplified with the LerayXT primers and samples amplified with the Foraminifera COI primers, and in the two treatments per dataset, *i.e.*, the sediment samples and the meiofauna samples. We calculated ASV richness of Foraminifera using the R package 'vegan' (v.2.5.7) (*Oksanen et al., 2020*).

Macher et al. (2022), *PeerJ*, DOI 10.7717/peerj.13952

6/27

Comparison of Foraminifera communities in sediment and meiofauna samples

For the comparison of Foraminifera community composition in sediment and meiofauna samples, we reduced the datasets to those 11 sampling sites for which we retained both sediment and meiofauna data after bioinformatic processing. We analysed the community composition using the 'adonis' PERMANOVA as implemented in the R package 'vegan', based on presence/absence of ASVs. The data was converted to presence/absence to avoid analyzing PCR biased data (*Elbrecht & Leese, 2015*; *Leray & Knowlton, 2017*). The factor 'sample treatment' (sediment sample vs. meiofauna sample) was used as the predictor and the Sørensen distances as response variable. Sørensen distances were calculated using the vegdist function implemented in the 'vegan' package. Communities were subsequently clustered with an average-linkage algorithm (hclust function) as in previous studies (*Burdon et al., 2016*; *Macher et al., 2018*). We regarded significant results with R^2 >0.09 (equivalent to $r = 0.30$) as moderate, and $R^2 > 0.25$ ($r = 0.50$) as strong (*Nakagawa & Cuthill, 2007*; *Cohen, 2013*). We visualised the number of ASVs found in both sediment and meiofauna samples and those found exclusively in one of the treatments with Venn diagrams using the R package 'VennDiagram' (v.1.7.1) (https://cran.r-project.org/package=VennDiagram).

Community composition of Foraminifera along the intertidal transect

The dataset used for analysis of the Foraminifera community along the intertidal transect comprised 18 sediment samples, *i.e.*, six samples along each of the three parallel transects spanning the intertidal zone from high tide to low tide line. We did not analyse the LerayXT dataset further because no Foraminifera were identified in this dataset. We also did not further analyse the meiofauna samples amplified with the Foraminifera primers, as amplification and sequencing success was low for these samples.

To assess whether we captured the whole richness of Foraminifera ASVs present in the study sites, we applied interpolation and extrapolation analyses as implemented in the R package iNEXT (v.2.0.20) (*Hsieh, Ma & Chao, 2016*). A TukeyHSD test as implemented in R was used to test for differences in ASV richness between samples at different tidal levels. Both iNEXT and Tukey HSD tests were performed based on previously published scripts (*Marwayana, Gold & Barber, 2021*). We analysed the community composition of Foraminifera ASVs along the intertidal transect using the 'adonis' PERMANOVA as described above, with 'transects' (one to three) and 'tidal level' as predictors, and Bray-Curtis distances as response variables. For the comparison of community composition in different tidal levels, we categorised the sampling sites into 'upper intertidal' (sampling locations S1, S2; six samples), 'middle intertidal' (S3, S4; six samples) and 'lower intertidal' (S5, S6; six samples). Non-Metric Multi-Dimensional Scaling (NMDS) plots were calculated and visualised with the R libraries 'vegan' and 'phyloseq' (*McMurdie & Holmes, 2013*).

RESULTS

Katwijk beach was classified as a tide modified beach with a reflective state based on beach morphology and the calculated Relative Tide Range (RTR) of 4.47 on the day of sampling.

See Fig. 1 for the beach profile and the location of sampling points along the three parallel transects.

Amplification and sequencing of taxonomic references

We applied single-cell metabarcoding to 36 specimens from 32 morphologically identified Foraminifera species. The quality check with the 'decontam' pipeline showed none of the ASVs to be a potential contamination. For 26 out of 32 sequenced species, we identified a single, species-specific ASV sequence. The exceptions were *Glabratellina* sp. (isolate number 17921) and *Planoglabratella opercularis* (isolate number 18053; both in the family Glabratellidae), which shared the same ASV sequence, and the same was found for *Trifarina earlandi* (isolate number 17160) and *Uvigerina bifurcata* (isolate number 17173; both in the family Uvigerinidae). This was confirmed by the additional Sanger sequenced specimens of *Trifarina*, *Uvigerina* and *Planoglabratella*, which show the same sequence as the ASV sequences obtained through single-cell metabarcoding. The two *Neoassilina ammonoides* specimens showed a distinct ASV sequence per specimen. See Supplementary Material 1 for ASV sequences and read numbers per sample.

The cladogram shows that the newly generated reference sequences cluster in the expected higher taxonomic groups (the non-monophyletic monothalamids, and the classes Globothalamea and Tubothalamea), and in the previously defined superfamilies and clades (*Pawlowski, Holzmann & Tyszka, 2013*; *Holzmann & Pawlowski, 2017*; *Holzmann et al., 2021*). However, *Cassidulinoides*, *Globocassidulina* and *Bolivina* (Serioidea) do not cluster with the other Serioidea, but as a sister clade to Glabratelloidea, the remaining Serioidea, Globigerinoidea and Rotaloidea. Further, *Stainforthia* and *Epistominella*, which are assigned to the globothalamid "Clade 3", do not cluster with the other "Clade 3" specimens, but with the Rotalioidea. See Fig. S1 for the cladogram.

The ASAP analysis delineated 50 lineages (ASAP score: 4.5, *P*-value: 2.67e–01, W: 3.96e–04) in the reference dataset and showed a best species delineation threshold of 0.82% genetic distance. Since previous ASAP analyses (*Girard et al., 2022*) showed that delineation thresholds can vary between taxonomic groups within Foraminifera, but are around 1% genetic distance, we further used 99% identity as the threshold for assigning sequences from the metabarcoding dataset to a species name.

Amplification and sequencing of community samples

Amplification of community samples was successful for 46 out of 48 samples with the LerayXT primers, and for 43 out of 48 samples using the Foraminifera COI primers. However, several meiofauna samples amplified with the Foraminifera COI primers as well as samples from above the high-tide line showed weak amplification, and these samples were subsequently lost during sequencing and bioinformatic processing. See below for details. The negative controls did not show any product (as tested by TapeStation with High Sensitivity Kit) and were therefore not sequenced. Sequencing resulted in 5,849,271 raw reads for the LerayXT sample, and 2,431,554 raw reads for the Foraminifera samples. After bioinformatic processing, the final LerayXT dataset comprised 2,963,461 merged and quality filtered sequences, and the Foraminifera dataset comprised 1,097,955 sequences. We

retained 43 samples for the LerayXT dataset (22 meiofauna samples, 21 sediment samples) containing 4025 ASVs, and 34 samples for the Foraminifera dataset (14 meiofauna samples and 20 sediment samples) containing 763 ASVs. See Supplementary Material 2 for read number per sample and nucleotide sequences.

Taxon- specificity of Foraminifera primers

Taxonomic annotation using Barcode of Life (BOLD) and the newly generated Foraminifera COI reference database showed that none of the reads in the LerayXT dataset were annotated to a Foraminifera reference sequence. In contrast, 176 ASVs out of the 763 ASVs were identified as Foraminifera in the samples amplified with the Foraminifera COI primers.

In the LerayXT-amplified meiofauna samples, most reads were annotated to Annelida (48%) and Arthropoda (28%), while 14% could not be assigned to a phylum. In contrast, 88.6% of reads in the Foraminifera meiofauna dataset were assigned to Foraminifera, and 9% could not be assigned to a taxonomic group. Other taxa (mostly: Annelida and Arthropoda) were present with less than 2% of reads. See Figs. 2A and 2B; and see Table S2 for a complete list of taxa.

In the LerayXT- amplified sediment samples, the largest fraction of reads (28%) could not be assigned to a phylum, 22% were assigned to Bacillariophyta (diatoms), and 19% to Arthropoda. In the sediment samples amplified with the Foraminifera primers, 88.9% of reads were annotated to a Foraminifera reference, and 9% were not assigned to a phylum. Other taxa were present with less than 2% of all reads. See Figs. 2C and 2D.

The majority of Foraminifera ASVs could not be identified on species level, but only on Class or Order level (see Table 1). The 138 Foraminifera ASVs that could be identified at least on class level belonged to the monothalamids, globothalamids and tubothalamids. See Table 2.

Comparison of Foraminifera communities in sediment and meiofauna samples

We retained 22 samples (11 meiofauna, 11 sediment) for the comparison of Foraminifera communities in sediment and meiofauna samples. These samples contained 154 foraminiferal ASVs. 71 ASVs (49.6%) were found in both sample types, 49 ASVs (34.3%) were exclusively found in the sediment samples, and 23 ASVs (16.1%) were only found in meiofauna samples (Fig. 3). The 'adonis' test implemented in vegan showed that ASV community composition differed significantly between meiofauna and sediment samples (R^2: 0.184, $p = 0.001 \ast\ast\ast$). ASVs assigned to Tubothalamea were found only in the meiofauna samples. Globothalamea ASVs were most common in meiofauna samples, while monothalamid ASVs were most common in sediment samples. See Table 3.

Foraminifera community composition along the intertidal transect

The final 'Foraminifera sediment' dataset for the analysis of community composition along the intertidal transect comprised 18 samples with 149 Foraminifera ASVs. The dataset comprised samples from six tidal levels (S1- high tide line, to S6- low tide line) with three biological replicates each. ASV richness did not differ significantly between tidal levels

Reads assigned to Phyla
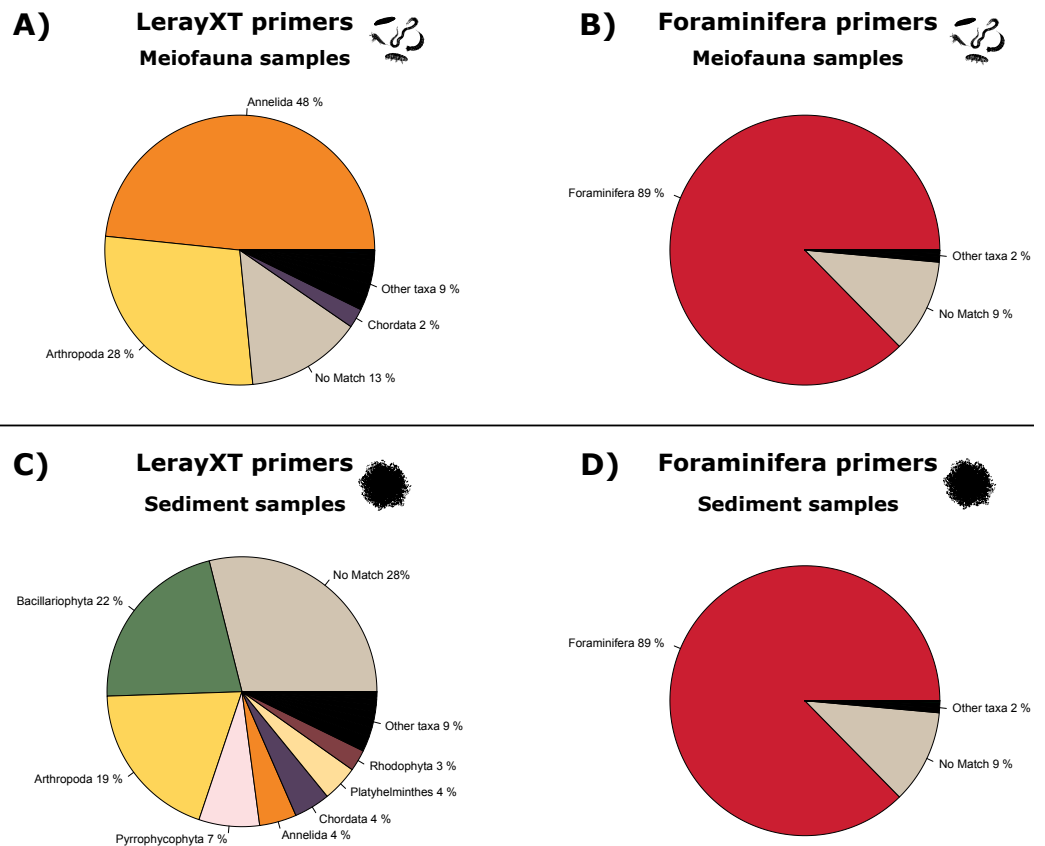
Figure 2 **Pie charts showing the proportion of reads annotated on phylum level.** (A) LerayXT meio-fauna samples, (B) Foraminifera meiofauna samples, (C) lerayXT sediment samples, (D) foraminifera sediment samples. Phyla present with <2% of annotated reads are summarised as ''other taxa' to improve readability.

Table 1 **Number and percentage of foraminifera ASVs identified on phylum, class, family, genus and species level.**

Taxonomic level	ASVs
Phylum level	176 (100%)
Class level	138 (78.41%)
Order level	116 (66.48%)
Family level	69 (39.2%)
Genus level	53 (30.11%)
Species level	9 (5.11%)

Table 2 Number and percentage of Foraminifera ASVs assigned to monothalamids, globothalemea and tubothalamea.

Taxonomic group	ASVs
Monothalamids	68 (49.28%)
Globothalamea	65 (47.1%)
Tubothalamea	5 (3.62%)

Number of Foraminifera ASVs in sediment and meiofauna samples
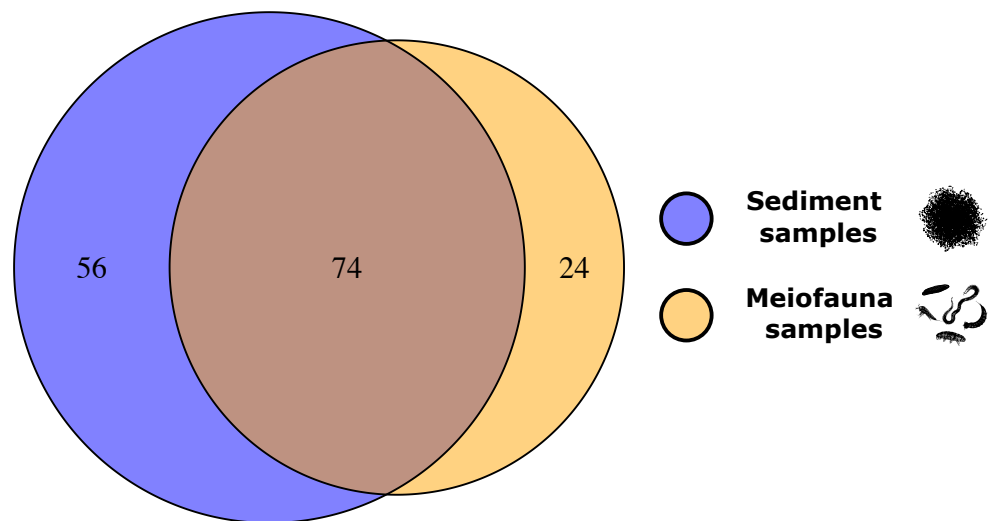
Figure 3 Venn diagrams showing Foraminifera ASVs exclusively found in either the sediment or meiofauna samples, and ASVs found with both sampling techniques.

Full-size 🖼 DOI: 10.7717/peerj.13952/fig-3

(TukeyHSD test results: Supplementary Material 3). Inter- and extrapolation showed that ASV richness approached saturation, with a sample coverage of 92.5%. See Fig. S2.

Analysis of community composition with 'adonis' PERMANOVA showed that Foraminifera ASV communities did not differ significantly between transects (R^2: 0.087, $p = 0.954$), but moderately between tidal levels (R^2: 0.187, $p = 0.005*$). NMDS plots (stress: 0.186) show that community composition in 'upper intertidal' samples differed the most from community composition in samples from the 'middle' and 'lower intertidal' area (Fig. 4). Communities in samples from the upper intertidal do not cluster together, showing heterogeneity of community composition, while samples from middle and lower intertidal cluster together. The majority of ASVs in the upper intertidal zone had a taxonomic match to monothalamids (59.81%). In the middle intertidal zone, 68.29% of ASVs were annotated to monothalamids, and in the lower intertidal zone, 75.29% of ASVs were annotated to monothalamids.

Table 3 Number and percentage of Foraminifera ASVs identified on class level in sediment and meiofauna samples.

Taxonomic Group	Meiofauna samples	Sediment samples
	No. of ASVs	No. of ASVs
Monothalamids	25 (31.25%)	55 (56.07%)
Globothalamea	50 (62.5%)	42 (43.3%)
Tubothalamea	5 (6.25%)	/
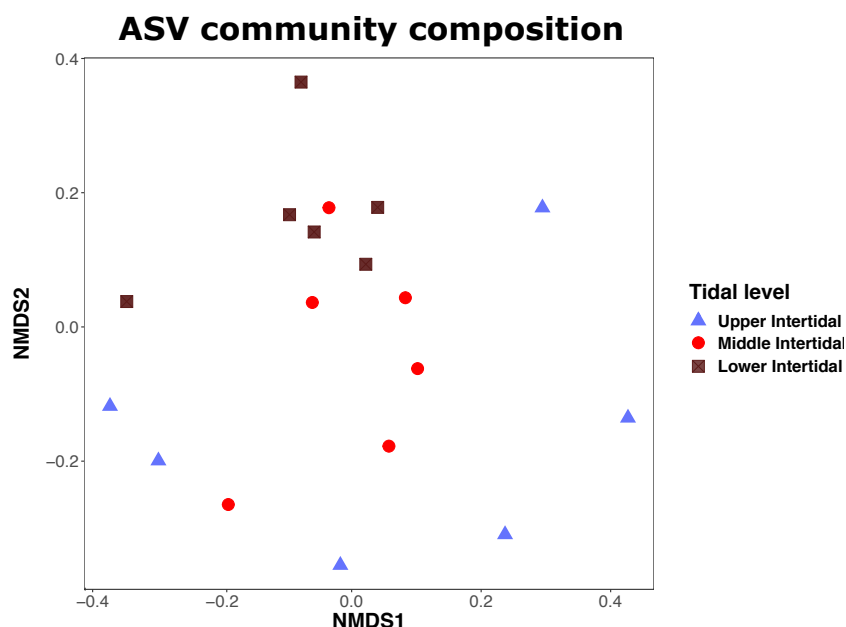
Figure 4 NMDS plot showing similarity of Foraminifera communities. NMDS plot showing similarity of Foraminifera communities (based on ASVs) between samples from upper, middle and lower intertidal sampling sites. Shapes and colours show the sampled tidal levels. Stress = 0.19.

Full-size ⬚ DOI: 10.7717/peerj.13952/fig-4

DISCUSSION

We report the first application of mitochondrial COI metabarcoding to infer Foraminifera communities from environmental and bulk samples, and show that the recently published Foraminifera COI primers amplify a wide range of foraminiferal taxa. The Foraminifera COI primers are highly taxon-specific, and we show that COI metabarcoding of Foraminifera can be used to infer changes in community composition along an intertidal transect. The availability of an easily obtainable molecular marker additionally to the commonly used 18S rRNA for metabarcoding of Foraminifera will be valuable for future studies, as it will open up new possibilities for studies on diversity and ecology of Foraminifera, and allows including multiple markers into molecular studies. This is becoming more common in metabarcoding studies (*Dupuis, Roe & Sperling, 2012*; *Gao et al., 2014*; *Fais et al., 2020*;

Eberle et al., 2020; *Gielings et al., 2021*) and could improve studies on phylogeny, species diversity and ecology of Foraminifera.

Amplification and sequencing of Foraminifera COI reference sequences

The calculated cladogram showed that all 49 newly generated reference sequences cluster in the expected taxonomic groups (monothalamids, and the classes Globothalamea and Tubothalamea). However, we point out that we cannot and do not want to make definite statements on the phylogeny of Foraminifera based on the analysed, short COI marker, since we calculated the cladogram as a quality check and are aware that short markers can result in weakly supported phylogenies. While the majority of species show a species-specific COI sequence, we note some exceptions where different species or genera share the same COI sequence, as seen in *Psammophaga* spp. (isolates 19260, 19296), *Glabratellina* sp. and *Planoglabratella opercularis,* and *Trifarina earlandi* and *Uvigerina bifurcata*. This indicates that the short COI marker used here does not always differentiate between closely related species or even genera. On the other hand, closely related species like the two *Rosalina* spp., which show distinct 18S rRNA based on previously published data, also show distinct COI sequences, and some species such as *Neoassilina ammonoides* show a high variability in their COI sequences. This has been documented before in Foraminifera (*Girard et al., 2022*), and similar discordance of morphological species identification and COI barcodes due to low variability of the barcoding region have been reported from a range of other taxonomic groups, *e.g.*, Amoebozoa (*Tekle, 2014*), arthropods (*Havemann et al., 2018*) and Porifera (*Yakhnenko & Itskovich, 2019*). Further research based on a higher number of single-cell (meta)barcoded specimens per well-identified morphospecies is therefore needed to investigate how reliable and consistently the COI metabarcoding marker can distinguish different closely related species or genera, and which species the marker might not resolve.

Amplification and sequencing of Foraminifera communities

We successfully amplified and sequenced DNA from intertidal meiofauna and sediment samples using both LerayXT and Foraminifera COI samples. However, amplification and sequencing success differed between primers, and between meiofauna and sediment samples. None of the reads in the LerayXT datasets was assigned to Foraminifera. While these primers were originally designed and are commonly used for metabarcoding metazoan taxa, they are known to amplify most main groups of Eukaryota and even prokaryotes (*Wangensteen et al., 2018*; *Garcés-Pastor et al., 2019*). The lack of Foraminifera sequences in this dataset therefore hints at a divergence of the primer binding sites in Foraminifera, hampering amplification. Similar results have been reported for primers targeting 18S rRNA, and subsequently Foraminifera-specific 18S primer combinations were developed (*de Vargas et al., 1997*; *Pawlowski, 2000*; *Morard et al., 2011*). When applying the Foraminifera COI primers, we retained 11 out of 24 meiofauna samples, and 20 out of 24 sediment samples. We speculate that the relatively low amplification and sequencing success observed for the meiofauna samples amplified with Foraminifera primers might

be due to the lower abundance of Foraminifera in these samples. This seems possible as previous studies on intertidal Foraminifera of the North Sea coast reported relatively low abundances of sometimes less than 10 ind/cm^3, (*Horton, Edwards & Lloyd, 1999*; *Müller-Navarra, Milker & Schmiedl, 2016*), while metazoan meiofaunal taxa like polychaetes can reach abundances of several thousand specimens per 10 cm^2 in North Sea beaches (*Kotwicki et al., 2005*). Thereby, the ratio of target DNA (Foraminifera) to non-target DNA (metazoan taxa) might have been unfavorable in the meiofauna samples. However, other studies on intertidal Foraminifera based on morphological analyses reported a high abundance, but also high variability between sites (*Reiter, 1959*; *Kameswara & Srinath, 2002*; *Lübbers & Schönfeld, 2018*), and higher abundances of Foraminifera are mostly reported from muddy sediments (*Morvan et al., 2006*; *Papaspyrou et al., 2013*). It is also possible that the MgCl$_2$ decantation method does not efficiently separate Foraminifera from sediment due to the calcareous test of many species, or that further amplification protocol optimisations like DNA dilutions are needed. In future studies, a size selection targeting smaller organisms, as used in previous studies on Foraminifera (*Peeters et al., 1999*; *Langezaal et al., 2003*; *Pawlowski et al., 2005*), could be applied to potentially increase the ratio of Foraminifera to metazoan meiofauna, and metabarcoding results should be compared to morphological analysis of samples.

Taxon-specificity of Foraminifera primers

We show that the majority of reads in both meiofauna and sediment samples amplified with the Foraminifera COI primers could be assigned to a foraminiferal reference, with about 90% of reads assigned to Foraminifera. A further 9% of reads could not be assigned to a taxonomic group. It seems possible that these reads are either foraminiferal taxa without available COI references, sequencing artifacts, random amplification of genomic regions, or nuclear mitochondrial DNA (NUMTs) (*Porter & Hajibabaei, 2021*; *Andújar et al., 2021*; *Graham, Gillespie & Krehenwinkel, 2021*). With growing COI reference for Foraminifera and understanding of their mitochondrial genomes, future studies could implement NUMT removal pipelines like METAMATE (*Andújar et al., 2021*) and defining thresholds for assigning COI sequences in metabarcoding datasets to taxonomic levels could be established. In this study, we applied ASAP delineation to the newly generated Foraminifera COI reference database and found that a threshold of about 1% genetic distance is appropriate for delineating species in Foraminifera COI datasets, which is in line with previous results (*Macher et al., 2021b*; *Girard et al., 2022*). However, more data on a wider range of species, genera and families is needed to define thresholds for higher taxonomic ranks, and we stress that the standard values we use in the present study, which stem from previous studies on eukaryotic COI, might actually over- or underestimate the number of taxa in Foraminifera metabarcoding datasets. Future studies should address this question by studying the diversity of foraminiferal COI, which could eventually lead to establishing a delineation system similar to that for 18S rRNA introduced by (*Morard et al., 2016*).

We also point out that even though our results show that the Foraminifera COI primers amplify a wide range of Foraminifera from all main taxonomic groups, it remains unknown

whether primers preferentially amplify certain Foraminifera taxa, thereby leading to primer bias as reported for other primers and taxonomic groups (*Tedersoo et al., 2015*; *Pawluczyk et al., 2015*; *Elbrecht & Leese, 2015*). Future studies should test this using mock communities of known taxonomic composition (*Smith et al., 2017*; *Lamb et al., 2019*), and by comparing COI marker to 18S rRNA marker data. Further, future studies should also test amplification of Foraminifera COI from other environments, *e.g.*, from plankton samples (*Morard et al., 2015*), soil (*Lejzerowicz et al., 2010*), mudflats (*Papaspyrou et al., 2013*) or tropical reefs (*Förderer, Rödder & Langer, 2018*) to test the amplification success for a wide range of sample types.

Comparison of Foraminifera communities in sediment and meiofauna samples

We found that the inferred community composition of Foraminifera differed significantly between meiofauna and sediment samples. Tubothalamea, albeit rare, were exclusively found in meiofauna samples. Monothalamids were more common in sediment samples. Previous studies based on 18S rRNA metabarcoding showed that sediment samples routinely contain a high number of monothalamids that can often not be reliably assigned to a lower taxonomic level (*Rodrigues et al., 2021*; *Nguyen et al., 2021*), which is in line with our results. We speculate that more small, amoeboid monothalamid taxa were lost during sieving and decantation of the meiofauna fraction, and that they were therefore present in higher abundance in the sediment samples. Corresponding to this, previous studies commonly reported differences in inferred community composition when targeting the same taxonomic groups with different sampling methods (*Brannock & Halanych, 2015*; *Macher et al., 2018*; *Holman et al., 2019*; *Castro et al., 2021*). Choosing the sampling and processing method should therefore be based on the targeted taxonomic group in future studies.

Foraminifera community composition along the intertidal transect

We show that the inferred community composition of Foraminifera in sediment samples changes along a transect from high tide line to low tide line. The Foraminifera communities in upper, middle and lower intertidal were dominated by monothalamids, but the fraction of monothalamids increased from the upper to the lower intertidal area. Since the available COI database of Foraminifera is still limited and the majority of ASVs could not be taxonomically assigned on family, genus or species level, more specific analyses on community composition are not possible until the number of COI reference barcodes increases. The increase in monothalamids towards the low-tide line might indicate that the environmental conditions (among others: sediment covered by seawater for longer time, more stable conditions due to lower wave energy, increased sediment grain size) might favor monothalamids over globothalameans and tubothalameans. Our findings are in line with previous studies based on morphology that show a change in intertidal Foraminifera communities corresponding to tidal levels (*Horton, 1999*; *Horton & Culver, 2008*; *Rush et al., 2021*), but we point out the need for more studies on Foraminifera COI metabarcoding to assess how meaningful ecological patterns can be extracted. Future studies

should also measure more parameters like sediment grain size, salinity, oxygen content and nutrient content of the sediment. Metabarcoding studies based on 18S rRNA have shown the versatility of the approach for studying foram community changes in different environments (*Pawlowski et al., 2014*; *Frontalini et al., 2018*; *Frontalini et al., 2020*; *Cordier et al., 2019*).

We cannot exclude that some of the found DNA is environmental DNA not usually live in the sandy beach ecosystem, as has been shown for many testate (*Taberlet et al., 2012*) or stems from Foraminifera that were washed ashore, but do Foraminifera (*Murray, 2006*). However, we point out that monothalamids, which were the most common taxa in our samples, have been regularly found in intertidal sandy beach ecosystems (*Larkin & Gooday, 2004*; *Golemansky, 2007*; *Alvarado & Goti, 2019*). Future studies should also compare the established 18S rRNA metabarcoding approach for Foraminifera to COI metabarcoding to assess how comparable the results are for inferring community composition. Furthermore, reference barcoding of species using both 18S and COI should be continued (*Pawlowski & Holzmann, 2014*; *Macher et al., 2021b*; *Girard et al., 2022*). To allow studying North Sea beach ecosystems in more detail, targeted barcoding of taxa known to live in intertidal habitats of Northern European coats should be performed, *e.g.*, more species of *Haynesina*, *Ammonia* and *Elphidium* (*Alve, 2001*; *Brouwer et al., 2015*). Since integrative approaches revealed a high diversity of yet unidentified Foraminifera and especially monothalamids (*Voltski & Pawlowski, 2015*), this taxonomic group needs special attention to fill gaps in reference libraries.

CONCLUSIONS

Our study shows that COI metabarcoding of foraminiferal communities with highly taxon-specific primers is feasible, thereby making mitochondrial metabarcoding available for further studies on Foraminifera. Continuing the build-up of Foraminifera COI reference databases based on morphologically identified species will be crucial to allow for species identification and refined ecological studies. Our results highlight that COI metabarcoding can be a powerful tool for assessing foraminiferal communities.

ACKNOWLEDGEMENTS

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Jan-Niklas Macher conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Dimitra Maria Bloska conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Maria Holzmann conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Elsa B. Girard conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Jan Pawlowski conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Willem Renema conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

All COI reference sequences are available in NCBI GenBank: OM719617–OM719669.

The metabarcoding raw data are available in NCBI Sequence Read Archive: PRJNA799248.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.13952#supplemental-information.

REFERENCES

Afgan E, Baker D, Batut B, vandenBeek M, Bouvier D, Cech M, Chilton J, Clements D, Coraor N, Grüning BA, Guerler A, Hillman-Jackson J, Hiltemann S, Jalili V, Rasche H, Soranzo N, Goecks J, Taylor J, Nekrutenko A, Blankenberg D. 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Research* **46**:W537–W544 DOI 10.1093/nar/gky379.

Macher et al. (2022), *PeerJ*, DOI 10.7717/peerj.13952

17/27

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology* **215**:403–410 DOI 10.1016/S0022-2836(05)80360-2.

Alvarado E, Goti I. 2019. Soft-shelled foraminifera from intertidal sandy beaches on the southeast of Azuero, Panama. *Visión Antataura* **3**:24–38.

Alve E. 2001. Temporal variability in vertical distributions of live (stained) intertidal foraminifera, Southern England. *The Journal of Foraminiferal Research* **31**:12–24 DOI 10.2113/0310012.

Andújar C, Creedy TJ, Arribas P, López H, Salces-Castellano A, Pérez-Delgado AJ, Vogler AP, Emerson BC. 2021. Validated removal of nuclear pseudogenes and sequencing artefacts from mitochondrial metabarcode data. *Molecular Ecology Resources* **21**:1772–1787 DOI 10.1111/1755-0998.13337.

Beentjes KK, Speksnijder AGC, Schilthuizen M, Hoogeveen M, Pastoor R, van der Hoorn BB. 2019. Increased performance of DNA metabarcoding of macroinverte-brates by taxonomic sorting. *PLOS ONE* **14**:e0226527 DOI 10.1371/journal.pone.0226527.

Borrelli C, Hou Y, Pawlowski JW, Holzmann M, Katz ME, Chandler GT, Bowser SS. 2018. Assessing SSU rDNA barcodes in Foraminifera: a case study using Bolivina quadrata. *The Journal of Eukaryotic Microbiology* **65**:220–235 DOI 10.1111/jeu.12471.

Brannock PM, Halanych KM. 2015. Meiofaunal community analysis by high-throughput sequencing: comparison of extraction, quality filtering, and clustering methods. *Marine Genomics* **23**:67–75 DOI 10.1016/j.margen.2015.05.007.

Brouwer MGM, Wolthers M, Hazeleger JH, Rossi F, Lourens LJ, Middelburg JJ, Duijnstee IAP. 2015. Differential response of intertidal foraminifera to community recovery following experimentally induced hypoxia. *Journal of Foraminiferal Research* **45**:220–234 DOI 10.2113/gsjfr.45.3.220.

Buchner D, Leese F. 2020. BOLDigger –a Python package to identify and organise sequences with the Barcode of Life Data systems. *Metabarcoding and Metagenomics* **4**:e53535 DOI 10.3897/mbmg.4.53535.

Burdon FJ, Reyes M, Alder AC, Joss A, Ort C, Räsänen K, Jokela J, Eggen RIL, Stamm C. 2016. Environmental context and magnitude of disturbance influence trait-mediated community responses to wastewater in streams. *Ecology and Evolution* **6**:3923–3939 DOI 10.1002/ece3.2165.

Burki F, Kudryavtsev A, Matz MV, Aglyamova GV, Bulman S, Fiers M, Keeling PJ, Pawlowski J. 2010. Evolution of Rhizaria: new insights from phyloge-nomic analysis of uncultivated protists. *BMC Evolutionary Biology* **10**:377 DOI 10.1186/1471-2148-10-377.

Castro LR, Meyer RS, Shapiro B, Shirazi S, Cutler S, Lagos AM, Quiroga SY. 2021. Metabarcoding meiofauna biodiversity assessment in four beaches of North-ern Colombia: effects of sampling protocols and primer choice. *Hydrobiologia* **848**:3407–3426 DOI 10.1007/s10750-021-04576-z.

Charnock MA, Jones RW. 1990. Agglutinated Foraminifera from the Palaeogene of the North Sea. *Paleoecology, Biostratigraphy, Paleoceanography and Taxonomy of Agglutinated Foraminifera* **327**:139–244 DOI 10.1007/978-94-011-3350-0_9.

Chronopoulou P-M, Salonen I, Bird C, Reichart G-J, Koho KA. 2019. Metabarcoding Insights into the trophic behavior and identity of intertidal benthic foraminifera. *Frontiers in Microbiology* **10**:1169 DOI 10.3389/fmicb.2019.01169.

Cohen J. 2013. *Statistical power analysis for the behavioral sciences.* New York: Routledge.

Collins RA, Bakker J, Wangensteen OS, Soto AZ, Corrigan L, Sims DW, Genner MJ, Mariani S. 2019. Non-specific amplification compromises environmental DNA metabarcoding with COI. *Methods in Ecology and Evolution* **10**:1985–2001 DOI 10.1111/2041-210x.13276.

Cordier T, Barrenechea I, Lejzerowicz F, Reo E, Pawlowski J. 2019. Benthic foraminiferal DNA metabarcodes significantly vary along a gradient from abyssal to hadal depths and between each side of the Kuril-Kamchatka trench. *Progress in Oceanography* **178**:102175 DOI 10.1016/j.pocean.2019.102175.

Darling KF, Kucera M, Wade CM. 2007. Global molecular phylogeography reveals persistent Arctic circumpolar isolation in a marine planktonic protist. *Proceedings of the National Academy of Sciences of the United States of America* **104**:5002–5007 DOI 10.1073/pnas.0700520104.

Darling KF, Wade CM. 2008. The genetic diversity of planktic foraminifera and the global distribution of ribosomal RNA genotypes. *Marine Micropaleontology* **67**:216–238 DOI 10.1016/j.marmicro.2008.01.009.

Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. 2018. Simple statistical identification and removal of contaminant sequences in marker-gene and metage-nomics data. *Microbiome* **6**:226 DOI 10.1186/s40168-018-0605-2.

de Vargas C, Zaninetti L, Hilbrecht H, Pawlowski J. 1997. Phylogeny and rates of molecular evolution of planktonic foraminifera: SSU rDNA sequences compared to the fossil record. *Journal of Molecular Evolution* **45(3)**:285–294 DOI 10.1007/PL00006232.

Dupuis JR, Roe AD, Sperling FAH. 2012. Multi-locus species delimitation in closely re-lated animals and fungi: one marker is not enough. *Molecular Ecology* **21**:4422–4436 DOI 10.1111/j.1365-294X.2012.05642.x.

Eberle J, Ahrens D, Mayer C, Niehuis O, Misof B. 2020. A plea for standardized nuclear markers in metazoan DNA taxonomy. *Trends in Ecology & Evolution* **35**:336–345 DOI 10.1016/j.tree.2019.12.003.

Edgar RC. 2016. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. DOI 10.1101/081257.

Elbrecht V, Leese F. 2015. Can DNA-based ecosystem assessments quantify species abun-dance? Testing primer bias and biomass-sequence relationships with an innovative metabarcoding protocol. *PLOS ONE* **10**:e0130324 DOI 10.1371/journal.pone.0130324.

Elbrecht V, Vamos EE, Meissner K, Aroviita J, Leese F. 2017. Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification

for routine stream monitoring. *Methods in Ecology and Evolution* **8**:1265–1275
DOI 10.1111/2041-210x.12789.

Fais M, Duarte S, Vieira PE, Sousa R, Hajibabaei M, Canchaya CA, Costa FO. 2020.
Small-scale spatial variation of meiofaunal communities in Lima estuary (NW
Portugal) assessed through metabarcoding. *Estuarine, Coastal and Shelf Science*
238:106683 DOI 10.1016/j.ecss.2020.106683.

Förderer M, Rödder D, Langer MR. 2018. Patterns of species richness and the center
of diversity in modern Indo-Pacific larger foraminifera. *Scientific Reports* **8**:8189
DOI 10.1038/s41598-018-26598-9.

Frontalini F, Cordier T, Balassi E, Armynotdu Chatelet E, Cermakova K, Apothéloz-
Perret-Gentil L, Martins MVA, Bucci C, Scantamburlo E, Treglia M, Bonamin V,
J Pawlowski. 2020. Benthic foraminiferal metabarcoding and morphology-based
assessment around three offshore gas platforms: congruence and complementarity.
Environment International **144**:106049 DOI 10.1016/j.envint.2020.106049.

Frontalini F, Greco M, Di Bella L, Lejzerowicz F, Reo E, Caruso A, Cosentino C,
Maccotta A, Scopelliti G, Nardelli MP, Losada MT, Châtelet EArmynotdu,
Coccioni R, Pawlowski J. 2018. Assessing the effect of mercury pollution on cultured
benthic foraminifera community using morphological and eDNA metabarcoding ap-
proaches. *Marine Pollution Bulletin* **129**:512–524 DOI 10.1016/j.marpolbul.2017.10.022.

Gao F, Gao S, Wang P, Katz LA, Song W. 2014. Phylogenetic analyses of cyclidiids
(Protista, Ciliophora, Scuticociliatia) based on multiple genes suggest their close
relationship with thigmotrichids. *Molecular Phylogenetics and Evolution* **75**:219–226
DOI 10.1016/j.ympev.2014.01.032.

Garcés-Pastor S, Wangensteen OS, Pérez-Haase A, Pèlachs A, Pérez-Obiol R, Cañellas
Boltà N, Mariani S, Vegas-Vilarrúbia T. 2019. DNA metabarcoding reveals modern
and past eukaryotic communities in a high-mountain peat bog system. *Journal of
Paleolimnology* **62**:425–441 DOI 10.1007/s10933-019-00097-x.

Gielings R, Fais M, Fontaneto D, Creer S, Costa FO, Renema W, Macher J-N. 2021.
DNA metabarcoding methods for the study of marine benthic meiofauna: a review.
Frontiers in Marine Science **8** DOI 10.3389/fmars.2021.730063.

Girard EB, Langerak A, Jompa J, Wangensteen OS, Macher J-N, Renema W.
2022. Mitochondrial cytochrome oxidase subunit 1: a promising molecular
marker for species identification in foraminifera. *Frontiers in Marine Science* **9**
DOI 10.3389/fmars.2022.809659.

Goldstein ST. 1999. Foraminifera: a biological overview. In: *Modern Foraminifera*.
Dordrecht: Springer, 37–55 DOI 10.1007/0-306-48104-9_3.

Golemansky V. 2007. Testate Amoebas and Monothalamous Foraminifera (Proto-
zoa) from the Bulgarian Black Sea Coast. *Biogeography and Ecology of Bulgaria*
82:555–570 DOI 10.1007/978-1-4020-5781-6_20.

Graham NR, Gillespie RG, Krehenwinkel H. 2021. Towards eradicating the nuisance of
numts and noise in molecular biodiversity assessment. *Molecular Ecology Resources*
21:1755–1758 DOI 10.1111/1755-0998.13414.

Greco M, Morard R, Kucera M. 2020. Single-cell metabarcoding reveals biotic interactions of the Arctic calcifier Neogloboquadrina pachyderma with the eukaryotic pelagic community. ArXiv preprint. arXiv:2020.10.20.347930 DOI 10.1101/2020.10.20.347930.

Gupta BKS. 2003. *Modern Foraminifera.* Dordrecht: Springer DOI 10.1007/0-306-48104-9.

Haenel Q, Holovachov O, Jondelius U, Sundberg P, Bourlat SJ. 2017. NGS-based biodiversity and community structure analysis of meiofaunal eukaryotes in shell sand from Hållö island, Smögen, and soft mud from Gullmarn Fjord, Sweden. *Biodiversity Data Journal* **5**:e12731.

Hallock P, Lidz BH, Cockey-Burkhard EM, Donnelly KB. 2003. Foraminifera as bioindicators in coral reef assessment and monitoring: the FORAM index, foraminifera in reef assessment and monitoring. *Environmental Monitoring and Assessment* **81**:221–238 DOI 10.1023/A:1021337310386.

Havemann N, Gossner MM, Hendrich L, Morinière J, Niedringhaus R, Schäfer P, Raupach MJ. 2018. From water striders to water bugs: the molecular diversity of aquatic Heteroptera (Gerromorpha, Nepomorpha) of Germany based on DNA barcodes. *PeerJ* **6**:e4577 DOI 10.7717/peerj.4577.

He X, Sutherland TF, Pawlowski J, Abbott CL. 2019. Responses of foraminifera communities to aquaculture-derived organic enrichment as revealed by environmental DNA metabarcoding. *Molecular Ecology* **28**:1138–1153 DOI 10.1111/mec.15007.

Holman LE, de Bruyn M, Creer S, Carvalho G, Robidart J, Rius M. 2019. Detection of introduced and resident marine species using environmental DNA metabarcoding of sediment and water. *Scientific Reports* **9**:11559 DOI 10.1038/s41598-019-47899-7.

Holovachov O. 2016. Metabarcoding of marine nematodes - evaluation of similarity scores used in alignment-based taxonomy assignment approach. *Biodiversity Data Journal* **4**:e10647.

Holzmann M, Gooday AJ, Siemensma F, Pawlowski J. 2021. Review: freshwater and soil foraminifera –a story of long-forgotten relatives. *Journal of Foraminiferal Research* **51**:318–331 DOI 10.2113/gsjfr.51.4.318.

Holzmann M, Hohenegger J, Apotheloz-Perret-Gentil L, Morard R, Abramovich S, Titelboim D, Pawlowski J. 2022. A revision of recent nummulitid genera based on molecular and morphological data reveals a new genus. *Journal of Earth Science* In press.

Holzmann M, Hohenegger J, Hallock P, Piller WE, Pawlowski J. 2001. Molecular phylogeny of large miliolid foraminifera (Soritacea Ehrenberg 1839). *Marine Micropaleontology* **43**:57–74 DOI 10.1016/s0377-8398(01)00021-4.

Holzmann M, Pawlowski J. 2017. An updated classification of rotaliid foraminifera based on ribosomal DNA phylogeny. *Marine Micropaleontology* **132**:18–34 DOI 10.1016/j.marmicro.2017.04.002.

Horton BP. 1999. The distribution of contemporary intertidal foraminifera at Cowpen Marsh, Tees Estuary, UK: implications for studies of Holocene sea-level changes. *Palaeogeography, Palaeoclimatology, Palaeoecology* **149**:127–149 DOI 10.1016/S0031-0182(98)00197-7.

Horton BP, Culver SJ. 2008. Modern intertidal foraminifera of the outer Banks, North Carolina, USA and their applicability for sea-level studies. *Journal of Coastal Research* **245**:1110–1125 DOI 10.2112/08a-0004.1.

Horton BP, Edwards RJ, Lloyd JM. 1999. UK intertidal foraminiferal distributions: implications for sea-level studies. *Marine Micropaleontology* **36**:205–223 DOI 10.1016/S0377-8398(99)00003-1.

Hsieh TC, Ma KH, Chao A. 2016. iNEXT: an R package for rarefaction and extrapolation of species diversity (H ill numbers). *Methods in Ecology and Evolution* **7**:1451–1456 DOI 10.1111/2041-210x.12613.

Kameswara RK, Srinath M. 2002. Foraminifera from beach sands along Saurashtra coast, north-west India. *Journal of the Marine Biological Association of India* **44**:22–36.

Katoh K, Misawa K, Kuma K-I, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* **30**:3059–3066 DOI 10.1093/nar/gkf436.

Keller G. 1983. Paleoclimatic analyses of middle Eocene through Oligocene planktic foraminiferal faunas. *Palaeogeography, Palaeoclimatology, Palaeoecology* **43**:73–94 DOI 10.1016/0031-0182(83)90049-4.

Kotwicki L, Troch MD, Urban-Malinga B, Gheskiere T, Węslawski JM. 2005. Horizontal and vertical distribution of meiofauna on sandy beaches of the North Sea (The Netherlands, Belgium, France). *Helgoland marine research* **59**:255–264 DOI 10.1007/s10152-005-0001-8.

Lamb PD, Hunter E, Pinnegar JK, Creer S, Davies RG, Taylor MI. 2019. How quantitative is metabarcoding: a meta-analytical approach. *Molecular Ecology* **28**:420–430 DOI 10.1111/mec.14920.

Langezaal AM, Ernst SR, Haese RR, van Bergen PF, van der Zwaan GJ. 2003. Disturbance of intertidal sediments: the response of bacteria and foraminifera. *Estuarine, Coastal and Shelf Science* **58**:249–264 DOI 10.1016/s0272-7714(03)00078-7.

Larkin KE, Gooday AJ. 2004. Soft-shelled monothalamous foraminifera at an intertidal site on the south coast of England. *Journal of Micropalaeontology* **23**:135–137 DOI 10.1144/jm.23.2.135.

Laroche O, Wood SA, Tremblay LA, Ellis JI, Lejzerowicz F, Pawlowski J, Lear G, Atalah J, Pochon X. 2016. First evaluation of foraminiferal metabarcoding for monitoring environmental impact from an offshore oil drilling site. *Marine Environmental Research* **120**:225–235 DOI 10.1016/j.marenvres.2016.08.009.

Lejzerowicz F, Pawlowski J, Fraissinet-Tachet L, Marmeisse R. 2010. Molecular evidence for widespread occurrence of Foraminifera in soils. *Environmental Microbiology* **12**:2518–2526 DOI 10.1111/j.1462-2920.2010.02225.x.

Leray M, Knowlton N. 2017. Random sampling causes the low reproducibility of rare eukaryotic OTUs in Illumina COI metabarcoding. *PeerJ* **5**:e3006 DOI 10.7717/peerj.3006.

Leray M, Yang JY, Meyer CP, Mills SC, Agudelo N, Ranwez V, Boehm JT, Machida RJ. 2013. A new versatile primer set targeting a short fragment of the mitochondrial COI

region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Frontiers in Zoology* **10**:34 DOI 10.1186/1742-9994-10-34.

Lübbers J, Schönfeld J. 2018. Recent saltmarsh foraminiferal assemblages from Iceland. *Estuarine, Coastal and Shelf Science* **200**:380–394 DOI 10.1016/j.ecss.2017.11.019.

Macher JN, van der Hoorn BB, Peijnenburg KTC, van Walraven L, Renema W. 2020. Metabarcoding reveals different zooplankton communities in northern and southern areas of the North Sea. ArXiv preprint. arXiv:2020.07.23.218479 DOI 10.1101/2020.07.23.218479.

Macher J-N, Prazeres M, Taudien S, Jompa J, Sadekov A, Renema W. 2021a. Integrating morphology and metagenomics to understand taxonomic variability of *Amphisorus* (Foraminifera, Miliolida) from Western Australia and Indonesia. *PLOS ONE* **16**:e0244616 DOI 10.1371/journal.pone.0244616.

Macher J-N, Vivancos A, Piggott JJ, Centeno FC, Matthaei CD, Leese F. 2018. Comparison of environmental DNA and bulk-sample metabarcoding using highly degenerate cytochrome c oxidase I primers. *Molecular Ecology Resources* **18**:1456–1468 DOI 10.1111/1755-0998.12940.

Macher J-N, Wideman JG, Girard EB, Langerak A, Duijm E, Jompa J, Sadekov A, Vos R, Wissels R, Renema W. 2021b. First report of mitochondrial COI in foraminifera and implications for DNA barcoding. *Scientific Reports* **11**:22165 DOI 10.1038/s41598-021-01589-5.

Magoč T, Salzberg SL. 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**:2957–2963 DOI 10.1093/bioinformatics/btr507.

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**:10 DOI 10.14806/ej.17.1.200.

Martínez A, Eckert EM, Artois T, Careddu G, Casu M, Curini-Galletti M, Gazale V, Gobert S, Ivanenko VN, Jondelius U, Marzano M, Pesole G, Zanello A, Todaro MA, Fontaneto D. 2020. Human access impacts biodiversity of microscopic animals in sandy beaches. *Communications Biology* **3**:175 DOI 10.1038/s42003-020-0912-6.

Marwayana ON, Gold Z, Barber PH. 2021. Environmental DNA in a global biodiversity hotspot: lessons from coral reef fish diversity across the Indonesian Archipelago. *Environmental DNA* **4**:222–238 DOI 10.1002/edn3.257.

McLachlan A, Defeo O, Short AD. 2018. Characterising sandy beaches into major types and states: implications for ecologists and managers. *Estuarine, Coastal and Shelf Science* **215**:152–160 DOI 10.1016/j.ecss.2018.09.027.

McMurdie PJ, Holmes S. 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLOS ONE* **8**:e61217 DOI 10.1371/journal.pone.0061217.

Morard R, Darling KF, Mahé F, Audic S, Ujiié Y, Weiner AKM, André A, Seears HA, Wade CM, Quillévéré F, Douady CJ, Escarguel G, deGaridel Thoron T, Siccha M, Kucera M, deVargas C. 2015. PFR2: a curated database of planktonic foraminifera 18S ribosomal DNA as a resource for studies of plankton ecology, biogeography and evolution. *Molecular Ecology Resources* **15**:1472–1485 DOI 10.1111/1755-0998.12410.

Morard R, Escarguel G, Weiner AKM, André A, Douady CJ, Wade CM, Darling KF, Ujiié Y, Seears HA, Quillévéré F, de Garidel-Thoron T, de Vargas C, Kucera M. 2016. Nomenclature for the nameless: a proposal for an integrative molecular taxonomy of cryptic diversity exemplified by planktonic foraminifera. *Systematic Biology* **65**:925–940 DOI 10.1093/sysbio/syw031.

Morard R, Quillevere F, Douady CJ, de Vargas C, de Garidel-Thoron T, Escarguel G. 2011. Worldwide genotyping in the planktonic foraminifer *Globoconella inflata*: implications for life history and paleoceanography. *PLOS ONE* **6(10)**:e26665 DOI 10.1371/journal.pone.0026665.

Morard R, Vollmar NM, Greco M, Kucera M. 2019. Unassigned diversity of planktonic foraminifera from environmental sequencing revealed as known but neglected species. *PLOS ONE* **14**:e0213936 DOI 10.1371/journal.pone.0213936.

Morvan J, Debenay JP, Jorissen F, Redois F, Bénéteau E, Delplancke M, Amato AS. 2006. Patchiness and life cycle of intertidal foraminifera: implication for environmental and paleoenvironmental interpretation. *Marine Micropaleontology* **61(1-3)**:131–154 DOI 10.1016/j.marmicro.2006.05.009.

Müller-Navarra K, Milker Y, Schmiedl G. 2016. Natural and anthropogenic influence on the distribution of salt marsh foraminifera in the Bay of Tümlau, German North Sea. *The Journal of Foraminiferal Research* **46**:61–74 DOI 10.2113/gsjfr.46.1.61.

Murray JW. 2006. Ecology and applications of Benthic Foraminifera DOI 10.1017/cbo9780511535529.

Nakagawa S, Cuthill IC. 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews of the Cambridge Philosophical Society* **82**:591–605 DOI 10.1111/j.1469-185X.2007.00027.x.

Nguyen N-L, Pawłowska J, Angeles IB, Zajaczkowski M, Pawłowski J. 2021. Metabarcoding reveals high diversity of benthic Foraminifera driven by atlantification of coastal Svalbard. *Research Square* DOI 10.21203/rs.3.rs-1009107/v1.

Nooijer LJ de, de Nooijer LJ, Duijnstee IAP, Bergman MJN, van der Zwaan GJ. 2008. The ecology of benthic foraminifera across the Frisian Front, southern North Sea. *Estuarine, Coastal and Shelf Science* **78**:715–726 DOI 10.1016/j.ecss.2008.02.015.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H, Oksanen MJ. 2020. vegan: community ecology package. R Package Version 2.5-7.

Papaspyrou S, Diz P, García-Robledo E, Corzo A, Jimenez-Arias JL. 2013. Benthic foraminiferal community changes and their relationship to environmental dynamics in intertidal muddy sediments (Bay of Cádiz, SW Spain). *Marine Ecology Progress Series* **490**:121–135 DOI 10.3354/meps10447.

Pawlowski J. 2000. Introduction to the molecular systematics of foraminifera. *Micropaleontology* **46**:1–12.

Pawlowski J, Esling P, Lejzerowicz F, Cedhagen T, Wilding TA. 2014. Environmental monitoring through protist next-generation sequencing metabarcoding: assessing the impact of fish farming on benthic foraminifera communities. *Molecular Ecology Resources* **14**:1129–1140 DOI 10.1111/1755-0998.12261.

Pawlowski J, Fahrni JF, Guiard J, Conlan K, Hardecker J, Habura A, Bowser SS. 2005. Allogromiid foraminifera and gromiids from under the Ross Ice Shelf: morphological and molecular diversity. *Polar Biology* **28**:514–522 DOI 10.1007/s00300-005-0717-6.

Pawlowski J, Holzmann M. 2002. Molecular phylogeny of Foraminifera a review. *European Journal of Protistology* **38**:1–10 DOI 10.1078/0932-4739-00857.

Pawlowski J, Holzmann M. 2014. A plea for DNA barcoding of Foraminifera. *Journal of Foraminiferal Research* **44**:62–67 DOI 10.2113/gsjfr.44.1.62.

Pawlowski J, Holzmann M, Tyszka J. 2013. New supraordinal classification of Foraminifera: molecules meet morphology. *Marine Micropaleontology* **100**:1–10 DOI 10.1016/j.marmicro.2013.04.002.

Pawlowski J, Lejzerowicz F, Esling P. 2014. Next-generation environmental diversity surveys of foraminifera: preparing the future. *The Biological Bulletin* **227**:93–106 DOI 10.1086/BBLv227n2p93.

Pawluczyk M, Weiss J, Links MG, Egaña Aranguren M, Wilkinson MD, Egea-Cortines M. 2015. Quantitative evaluation of bias in PCR amplification and next-generation sequencing derived from metabarcoding samples. *Analytical and Bioanalytical Chemistry* **407**:1841–1848 DOI 10.1007/s00216-014-8435-y.

Peeters F, Ivanova E, Conan S, Brummer G-J, Ganssen G, Troelstra S, van Hinte J. 1999. A size analysis of planktic foraminifera from the Arabian Sea. *Marine Micropaleontology* **36**:31–63 DOI 10.1016/s0377-8398(98)00026-7.

Porter TM, Hajibabaei M. 2021. Profile hidden Markov model sequence analysis can help remove putative pseudogenes from DNA barcoding and metabarcoding datasets. *BMC Bioinformatics* **22**:256 DOI 10.1186/s12859-021-04180-x.

Prazeres M, Morard R, Roberts TE, Doo SS, Jompa J, Schmidt C, Stuhr M, Renema W, Kucera M. 2020. High dispersal capacity and biogeographic breaks shape the genetic diversity of a globally distributed reef-dwelling calcifier. *Ecology and Evolution* **10**:5976–5989 DOI 10.1002/ece3.6335.

Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 –approximately maximum-likelihood trees for large alignments. *PLOS ONE* **5**:e9490 DOI 10.1371/journal.pone.0009490.

Puillandre N, Brouillet S, Achaz G. 2021. ASAP: assemble species by automatic partitioning. *Molecular Ecology Resources* **21**:609–620 DOI 10.1111/1755-0998.13281.

Ratnasingham S, Hebert PDN. 2007. bold: the barcode of life data system (http://www.barcodinglife.org). *Molecular Ecology Notes* **7**:355–364 DOI 10.1111/j.1471-8286.2007.01678.x.

Reiter M. 1959. Seasonal variations in Intertidal Foraminifera of Santa Monica Bay, California. *Journal of Paleontology* **33**:606–630.

Renema W. 2018. Terrestrial influence as a key driver of spatial variability in large benthic foraminiferal assemblage composition in the Central Indo-Pacific. *Earth-Science Reviews* **177**:514–544 DOI 10.1016/j.earscirev.2017.12.013.

Rodrigues AR, Pellizari VH, de Salles Almeida C, de Sousa ECPM, Holzmann M, Pawlowski J. 2021. New phylotypes of foraminifera in subtropical Brazilian coastal

waters revealed by environmental DNA metabarcoding. *Journal of Sedimentary Environments* **6**:13–23 DOI 10.1007/s43217-021-00049-4.

Rush G, McDarby P, Edwards R, Milker Y, Roland Garrett E, Gehrels W. 2021. Development of an intertidal foraminifera training set for the North Sea and an assessment of its application for Holocene sea-level reconstructions. *Marine Micropaleontology* **169**:102055 DOI 10.1016/j.marmicro.2021.102055.

Scheibner C, Speijer RP, Marzouk AM. 2005. Turnover of larger foraminifera during the Paleocene-Eocene Thermal Maximum and paleoclimatic control on the evolution of platform ecosystems. *Geology* **33**:493 DOI 10.1130/g21237.1.

Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**:863–864 DOI 10.1093/bioinformatics/btr026.

Schnell IB, Bohmann K, Gilbert MTP. 2015. Tag jumps illuminated–reducing sequence-to-sample misidentifications in metabarcoding studies. *Molecular Ecology Resources* **15**:1289–1303 DOI 10.1111/1755-0998.12402.

Schweizer M, Schweizer Magali, Fontaine D, Pawlowski J. 2011. Phylogenetic position of two Patagonian Cibicididae (Rotaliida, Foraminifera): Cibicidoides dispars (d'Orbigny, 1839) and Cibicidoides variabilis (d'Orbigny, 1826). *Revue de Micropaléontologie* **54**:175–182 DOI 10.1016/j.revmic.2011.03.002.

Siemensma F, Apothéloz-Perret-Gentil L, Holzmann M, Clauss S, Völcker E, Pawlowski J. 2017. Taxonomic revision of freshwater foraminifera with the description of two new agglutinated species and genera. *European Journal of Protistology* **60**:28–44 DOI 10.1016/j.ejop.2017.05.006.

Smith KF, Kohli GS, Murray SA, Rhodes LL. 2017. Assessment of the metabarcoding approach for community analysis of benthic-epiphytic dinoflagellates using mock communities. *New Zealand Journal of Marine and Freshwater Research* **51**:555–576 DOI 10.1080/00288330.2017.1298632.

Somerfield PJ, Warwick RM, Moens T. 2005. Meiofauna techniques. In: *Methods for the study of marine benthos.* Chichester: John Wiley & Sons, 229–272.

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

Tamura K, Stecher G, Kumar S. 2021. MEGA11: molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution* **38**:3022–3027 DOI 10.1093/molbev/msab120.

Tedersoo L, Anslan S, Bahram M, Põlme S, Riit T, Liiv I, Kõljalg U, Kisand V, Nilsson H, Hildebrand F, Bork P, Abarenkov K. 2015. Shotgun metagenomes and multiple primer pair-barcode combinations of amplicons reveal biases in metabarcoding analyses of fungi. *MycoKeys* **10**:1–43 DOI 10.3897/mycokeys.10.4852.

Tekle YI. 2014. DNA barcoding in amoebozoa and challenges: the example of Cochliopodium. *Protist* **165**:473–484 DOI 10.1016/j.protis.2014.05.002.

Turon X, Antich A, Palacín C, Præbel K, Wangensteen OS. 2019. From metabarcoding to metaphylogeography: separating the wheat from the chaff. DOI 10.1101/629535.

Voltski I, Pawlowski J. 2015. Flexammina islandica gen, nov. sp. nov. and some new phylotypes of monothalamous foraminifera from the coast of Iceland. *Zootaxa* **3964**:245–259 DOI 10.11646/zootaxa.3964.2.5.

Wangensteen OS, Palacín C, Guardiola M, Turon X. 2018. DNA metabarcoding of littoral hard-bottom communities: high diversity and database gaps revealed by two molecular markers. *PeerJ* **6**:e4705 DOI 10.7717/peerj.4705.

Weber AA-T, Pawlowski J. 2014. Wide occurrence of SSU rDNA intragenomic polymorphism in foraminifera and its implications for molecular species identification. *Protist* **165**:645–661 DOI 10.1016/j.protis.2014.07.006.

Weigand AM, Macher J-N. 2018. A DNA metabarcoding protocol for hyporheic freshwater meiofauna: evaluating highly degenerate COI primers and replication strategy. *Metabarcoding and Metagenomics* **2**:e26869 DOI 10.3897/mbmg.2.26869.

Yakhnenko AS, Itskovich VB. 2019. Analysis of mtDNA variability in closely related Baikal sponge species for new barcoding marker development. *Limnology/The Japanese Society of Limnology* **21**:49–57.