scientific reports

Check for updates

OPEN

Integrating DNA metabarcoding and morphological analysis improves marine zooplankton biodiversity assessment

So-yeon Kim[1,5], Seongmin Cheon[2,5], Chungoo Park[1,2,3]✉ & Ho Young Soh[1,4]✉

Marine copepod communities play crucial roles in ocean ecosystems. However, their accurate assessment remains challenging due to taxonomic complexities. This study combines morphological and DNA metabarcoding approaches to evaluate copepod diversity and community structure in the northern East China Sea. Zooplankton samples were collected from 10 stations along a coastal-offshore gradient in August 2019. Morphological analysis identified 34 species from 25 genera, while DNA metabarcoding detected 31 species from 20 genera. Both methods revealed distinct coastal and offshore assemblages, with *Paracalanus parvus* s.l. as the dominant species across all stations. A significant positive correlation was found between morphology-based individual counts and metabarcoding sequence reads (Spearman's *Rho* = 0.58, $p < 0.001$), improving at the genus level (*Rho* = 0.70, $p < 0.001$). Redundancy analysis revealed that salinity, temperature, and phytoplankton density significantly influenced copepod distribution. Although both approaches captured similar broad-scale patterns, they provided complementary insights into community structure. Morphological identification was more effective for detecting Cyclopoida diversity, whereas DNA metabarcoding had greater sensitivity for specific Calanoid species. This study underscores the value of integrating traditional and molecular methods for marine biodiversity assessment, especially in the context of global environmental changes.

Keywords Copepod diversity, DNA metabarcoding, Morphological identification, East China Sea, Zooplankton ecology, Marine biodiversity assessment

Abbreviations

nECS	Northern East China Sea
CDW	Changjiang diluted water
TWC	Taiwan warm current
YSCW	Yellow sea cold water
CTD	Conductivity-temperature-depth
Chl-a	Chlorophyll-a
gDNA	Genome DNA
COI	Cytochrome c oxidase subunit 1
taxID	Taxonomic identifier code
NT	Non-redundant nucleotide sequence
mOTU	Molecular operational taxonomic unit
clr	Centered-log ratio
eDNA	Environmental DNA
nMDS	Non-metric multi-dimensional scaling
RDA	Redundancy analysis
T-S	Temperature-salinity

[1]Interdisciplinary Program of Bigdata Fishery Resources Management, Chonnam National University, Yeosu, Republic of Korea. [2]School of Biological Sciences and Technology, Chonnam National University, Yongbong-ro 77, Gwangju 61186, Republic of Korea. [3]Institute of Systems Biology and Life Science Informatics, Chonnam National University, Yongbong-ro 77, Gwangju 61186, Republic of Korea. [4]Department of Ocean Integrated Science, Chonnam National University, Yeosu, Republic of Korea. [5]These authors contributed equally: So-yeon Kim and Seongmin Cheon. ✉email: chungoo@jnu.ac.kr; hysoh@jnu.ac.kr

Marine ecosystems are complex and dynamic, with zooplankton communities playing critical roles in their structure and function. Among zooplankton, copepods are particularly significant, comprising approximately 70% of marine zooplankton and serving as key links in oceanic food webs[1]. With over 14,946 described species worldwide (World Register of Marine Species, 2024), copepods exhibit remarkable phylogenetic and functional diversity, occupying a wide range of aquatic niches from benthic to planktonic environments[2–9].

The northern East China Sea (nECS) offers a unique and challenging environment for studying copepod communities. As a marginal sea where land and ocean converge, the nECS is characterized by intricate current patterns and diverse water masses. Several key hydrographic features influence this region. The Changjiang Diluted Water (CDW) brings cold, fresh, nutrient-rich water from the Yangtze River during the summer months[10]. At the same time, the Taiwan Warm Current (TWC) transports warm, tropical waters from the northeast of Taiwan[11]. The Yellow Sea Cold Water (YSCW) further contributes to the complexity, creating stable stratification with warm upper layers and cold lower layers. These hydrodynamic conditions form a mosaic of habitats that support higher copepod diversity in offshore areas compared to nearshore environments. Understanding copepod distribution and community structure in this region is essential for assessing marine ecosystem health and predicting responses to environmental changes.

However, accurately assessing copepod diversity and community composition poses significant challenges. Traditional morphological identification methods, though directly linked to established taxonomic knowledge, are labor-intensive and require considerable expertise. These methods are especially challenging for small copepods (1–5 mm) and are complicated by cryptic speciation[12–14]. Additionally, the decline in trained taxonomists raises concerns about the consistency and accuracy of morphological identification across studies[12]. In recent years, DNA metabarcoding has emerged as a powerful complementary approach to traditional morphological methods. This molecular technique enables high-throughput analysis of bulk samples, potentially capturing cryptic diversity and improving the identification of immature stages[15]. However, metabarcoding also has limitations, including primer bias, the need for comprehensive reference databases, and challenges in the quantitative interpretation of sequence data[16,17].

Significant advances in zooplankton metabarcoding methodology over the past decade have substantially enhanced our understanding of marine biodiversity. Earlier studies established foundation protocols for analyzing zooplankton communities and revealed previously undetected diversity in marine ecosystems[14,18]. This work was followed by methodological refinements that improved taxonomic coverage through refined primer design and addressing quantification challenges with enhanced abundance estimation protocols[19,20]. Recent developments in multi-marker approaches have further expanded our capacity for comprehensive biodiversity assessments[21]. However, despite these methodological advances, most studies have focused on either morphological or molecular approaches in isolation, limiting our understanding of their complementary strengths and weaknesses in different marine environments.

To address these methodological challenges in copepod diversity assessment, this study employed an integrated approach combining morphological analysis and DNA metabarcoding using the COI gene marker. We specifically aimed to (1) compare the effectiveness of morphological and molecular methods in capturing copepod diversity across the coastal-offshore gradient in the nECS, (2) evaluate the quantitative relationship between traditional abundance estimates and metabarcoding read numbers, and (3) assess how environmental factors influence copepod community structure. This integrated assessment not only enhances our understanding of marine biodiversity dynamics but also provides insights into the complementarity of traditional and molecular approaches for marine ecosystem monitoring.

Results

Hydrographic structure of the northern East China Sea

Oceanographic surveys conducted at 21 stations in the nECS from August 5 to 8, 2019, revealed distinct spatial patterns in temperature and salinity (Fig. 1, Supplementary Table S1). Surface water temperatures ranged from 15.22 to 25.13 °C (mean: 20.06 ± 2.95 °C, Supplementary Fig. S1a–c), with the highest temperatures recorded near the Changjiang River mouth (stations EC14, EC16; mean: 29.16 ± 0.15 °C). Offshore stations (EC02, EC04, EC06, EC08, EC10, EC12, EC18, EC20) exhibited slightly lower surface temperatures, with a mean of 28.19 ± 0.29 °C. Vertical temperature profiles showed strong stratification at most stations (Supplementary Fig. S1a–c), typically comprising a warm surface layer (0–20 m), a thermocline (20–50 m), and a cold bottom layer (> 50 m).

Salinity distributions exhibited a clear coastal-offshore gradient (Supplementary Fig. S1d–f). Lower salinities (mean: 29.47 ± 0.61 PSU) were recorded near the Changjiang River and Chinese coastal regions (EC02, EC12, EC14, EC16, EC18), while higher salinities (mean: 31.43 ± 0.71 PSU) were observed at offshore stations (EC04, EC06, EC08, EC10, EC20), particularly at depths greater than 100 m. Salinities exceeding 34 PSU were detected only at stations EC06 and EC08.

Temperature-salinity (T-S) analysis at 10 representative stations (EC02, EC04, EC06, EC08, EC10, EC12, EC14, EC16, EC18, EC20; Fig. 2) revealed three distinct water masses: Changjiang Diluted Water (CDW), Taiwan Warm Current (TWC), and Yellow Sea Cold Water (YSCW). The CDW (salinity < 31 PSU, temperature > 20 °C) was prominent near the Changjiang River mouth (EC02, EC12, EC14, EC16). The TWC (salinity > 33 PSU, temperature > 20 °C) was primarily observed at offshore stations (EC06, EC08, EC10, EC20). The YSCW (temperature < 10 °C, salinity 32–33 PSU) was detected in bottom layers at stations deeper than 50 m (EC04, EC06, EC08, EC10, EC20). These 10 stations, encompassing diverse environmental gradients, were selected for both morphological and DNA metabarcoding analyses of zooplankton communities.

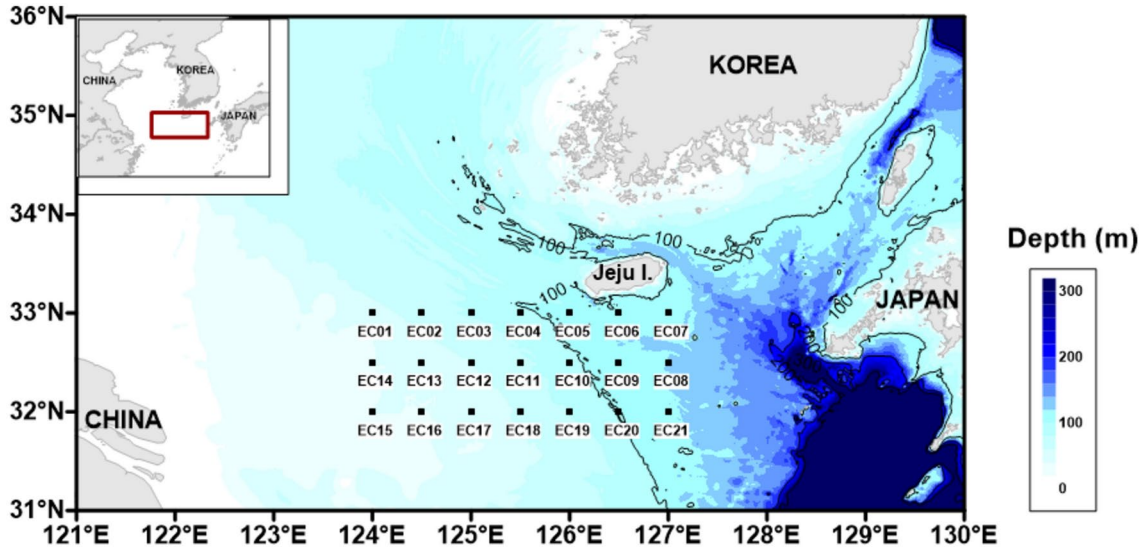


Fig. 1. Sampling stations in the northern East China Sea. Black dots indicate the 21 sampling stations (EC01–EC21). Red squares mark the 10 stations (EC02, EC04, EC06, EC08, EC10, EC12, EC14, EC16, EC18, EC20) selected for both morphological and DNA metabarcoding analyses. Bathymetry is illustrated with blue shading, with 100 m depth contours.

Taxonomic composition of copepods based on morphological analysis

Morphological analysis revealed a diverse copepod community consisting of 34 species from 25 genera, representing two orders: Calanoida and Cyclopoida (Fig. 3). The community was dominated by four genera: *Calanus*, *Paracalanus*, *Oithona*, and *Ditrichocorycaeus*. A clear spatial pattern in copepod diversity was observed, with the eastern stations (EC06, EC20, EC08) showing the highest species richness. Station EC06, in particular, emerged as a biodiversity hotspot, hosting 23 species (68% of the total), including 18 oceanic warm-water species and five neritic species.

At the species level, *Paracalanus parvus* s.l. was the most dominant and widespread species, present at all stations, with abundances ranging from 434 to 3341 individuals/m^3. *Calanus sinicus* displayed a broad distribution, with peak abundances recorded at stations EC16 (423 individuals/m^3) and EC02 (135 individuals/m^3). *Acartia pacifica* was predominantly found in coastal waters, while *Pleuromamma* species were restricted to offshore stations. At the genus level, *Oithona* was widely distributed, with *Oithona similis* particularly abundant at stations EC10, EC20, and EC04. *Temora* displayed a preference for coastal areas, while *Clausocalanus* and *Euchaeta* were mainly observed at stations influenced by the TWC. Family-level analysis revealed that Paracalanidae was the most widespread family, followed by Calanidae. Acartiidae showed a clear preference for coastal and transitional waters, whereas Euchaetidae and Metridinidae were largely restricted to offshore stations.

Taxonomic composition of copepods based on DNA metabarcoding analysis

DNA metabarcoding identified 31 species from 20 genera, belonging to 15 families and two orders (Fig. 4). The order Calanoida dominated, comprising 13 families, 18 genera, and 29 species, while Cyclopoida was represented by two families, two genera, and two species. Station EC06 had the highest species richness, with 21 species, followed by stations EC08 and EC20, which each recorded 18 species. In contrast, stations EC18 and EC12 exhibited the lowest diversity, with six and seven species, respectively.

Paracalanus parvus s.l. emerged as the most dominant and widespread species, present at all stations, with sequence reads ranging from 252,233 to 2,383,129 × 10^6/m^3. Other widely distributed species included *Acartia pacifica*, *Calanus sinicus*, *Temora discaudata*, *Oithona similis*, and *Ditrichocorycaeus affinis*. Several species showed clear habitat preferences: *Acartia bifilosa* was exclusively found in coastal waters (EC02), while *Pareucalanus attenuatus* was detected only in offshore waters (EC20). *Lucicutia flavicornis* and *Calocalanus plumulosus* were exclusively found at station EC06. At the genus level, *Acartia*, *Calanus*, *Paracalanus*, and *Euchaeta* were the most frequently detected. Family-level analysis supported the morphological findings, with Paracalanidae, Calanidae, and Acartiidae showing widespread distributions across the sampled stations.

Comparative analysis of copepod diversity: integrating morphological and DNA metabarcoding approaches

The concordance between morphological and DNA metabarcoding approaches decreased from higher to lower taxonomic levels (Fig. 5). At the family level, 70% of families were detected by both methods. Genus-level agreement was 60%, while species-level agreement dropped to 50%. Order-specific analysis revealed distinct patterns for Calanoida and Cyclopoida. In Calanoida, DNA metabarcoding slightly outperformed morphological

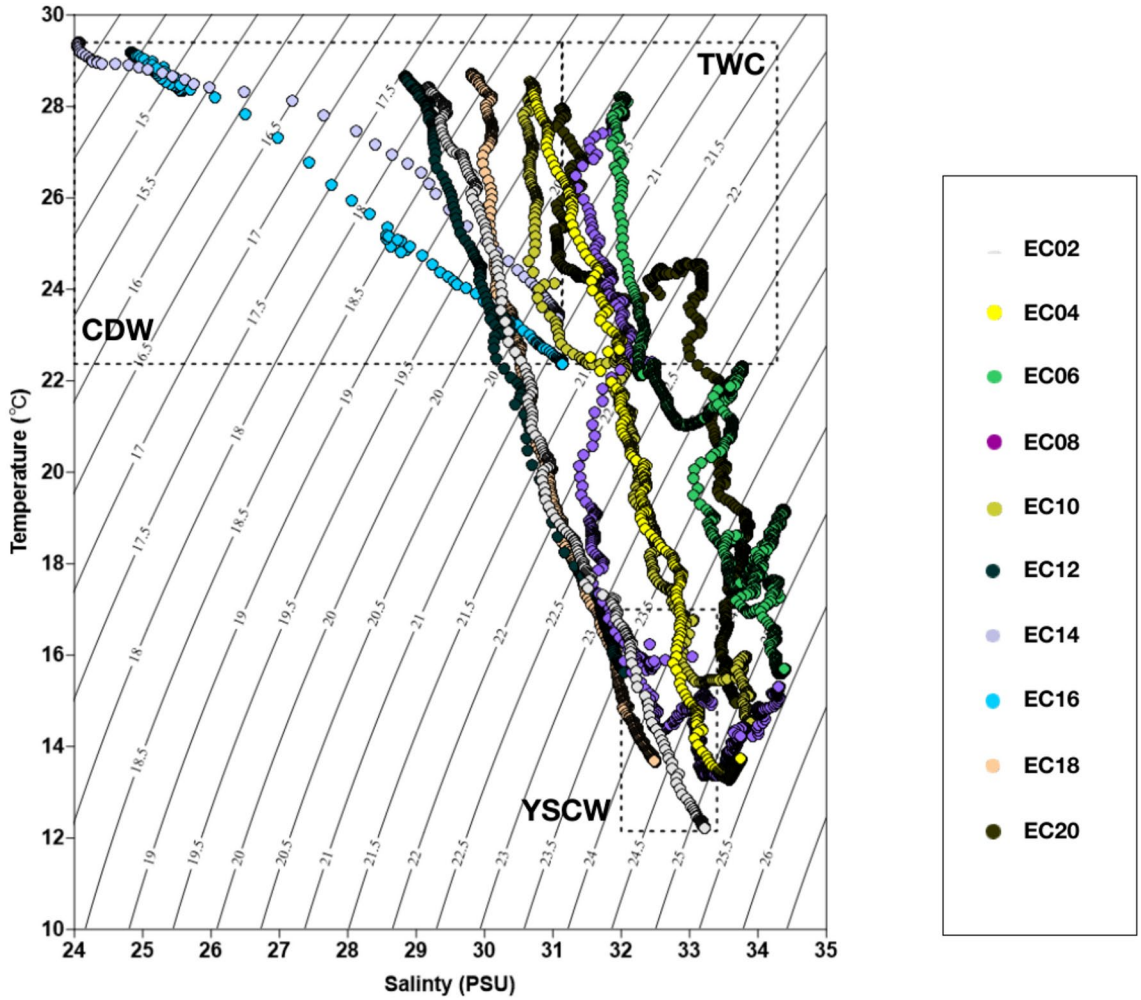


Fig. 2. Temperature-Salinity (T-S) diagram of the northern East China Sea. Water masses are labeled as follows: CDW (Changjiang Diluted Water), TWC (Taiwan Warm Current), and YSCW (Yellow Sea Cold Water). Colored lines represent the T-S profiles from the 10 sampling stations (EC02–EC20).

analysis in species detection. However, for Cyclopoida, morphological analysis identified significantly more species than DNA metabarcoding.

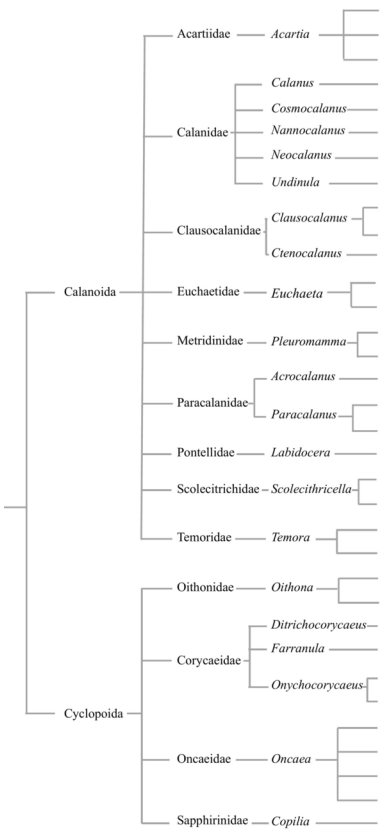
The quantitative relationship between two methods were further examined using normalized metabarcoding data (as described in Materials and Methods). Correlation analysis between morphology-based individual counts and DNA metabarcoding sequence reads for 16 common species showed a significant positive correlation (Spearman's $Rho = 0.58$, $p = 1.32 \times 10^{-6}$) (Fig. 6). This correlation demonstrates that both methods captured similar distribution patterns of species abundance across sampling stations. Similar abundance patterns were also observed in each of the 16 common species across different stations (Supplementary Fig. S2). To account for potential biases due to immature stages (copepodites) in morphological identification, we conducted an additional analysis at the genus level. This genus-level analysis revealed a stronger positive correlation (Spearman's $Rho = 0.70$, $p = 6.89 \times 10^{-10}$) compared to the species-level analysis (Supplementary Fig. S3 and S4).

Community structure analysis using Bray–Curtis dissimilarity revealed two distinct groups in both methods. Hierarchical clustering and nMDS analyses supported the division of stations into two main groups, broadly corresponding to coastal (Group A) and offshore (Group B) stations (Fig. 7 and Supplementary Fig. S5).

Redundancy analysis (RDA) was conducted to examine the relationship between copepod species distribution and environmental factors. Environmental parameters (temperature, salinity, turbidity, and densities of diatoms, dinoflagellates, and ciliates) and zooplankton samples were analyzed from the same depth range (surface to approximately 10 m), ensuring direct correspondence in the analysis. In the morphological data, the first two RDA axes explained 63.87% of the total variation in copepod assemblages. The DNA metabarcoding data showed similar patterns, with the first two RDA axes accounting for 61.51% of the total variation (Fig. 8).

Discussion

Morphological identification provided a detailed overview of the copepod community structure across the nECS, identifying 34 species from 25 genera. This method offers the advantage of directly linking observed



	EC02	EC04	EC06	EC08	EC10	EC12	EC14	EC16	EC18	EC20
Acartia (Acartia) danae	-	-	25	-	-	-	-	-	-	45
Acartia (Acartiura) omorii	-	23	-	-	-	-	-	-	-	-
Acartia (Odontacartia) pacifica	17	-	-	-	-	19	22	212	18	15
Calanus sinicus	135	-	15	31	31	38	-	423	18	15
Cosmocalanus darwinii	-	-	-	-	-	-	-	-	-	15
Nannocalanus minor	-	-	5	-	-	-	-	-	-	-
Neocalanus robustior	-	-	20	-	-	-	-	-	-	-
Undinula vulgaris	-	-	10	7	-	-	-	-	-	15
Clausocalanus farrani	-	-	25	-	-	-	-	-	-	-
Clausocalanus furcatus	-	-	53	-	25	-	-	-	-	-
Ctenocalanus vanus	-	-	49	-	13	-	-	-	-	45
Euchaeta concinna	-	-	29	19	13	-	-	-	-	15
Euchaeta plana	-	-	5	43	-	-	-	-	-	-
Pleuromamma abdominalis	-	-	-	7	-	-	-	-	-	60
Pleuromamma gracilis	-	-	-	-	-	-	-	-	-	60
Acrocalanus gibber	-	-	-	-	-	-	-	-	-	15
Paracalanus aculeatus	-	-	10	-	-	-	22	136	-	15
Paracalanus parvus s. l.	3,341	2,975	972	941	467	434	777	1,480	657	2,360
Labidocera acuta	-	-	-	7	-	-	-	-	-	-
Scolecithricella longispinosa	-	-	10	-	-	-	-	-	-	15
Scolecithricella nicobarica	-	-	-	7	-	-	-	-	-	-
Temora discaudata	-	-	5	-	-	-	22	16	-	-
Temora turbinata	-	-	-	-	-	5	87	76	9	-
Oithona fallax	17	-	241	97	13	-	-	-	18	104
Oithona similis	-	270	106	211	522	19	-	76	99	297
Ditrichocorycaeus affinis	-	23	20	13	-	5	518	76	18	15
Farranula gibbula	-	-	-	13	-	-	-	-	-	-
Onychocorycaeus agilis	-	-	5	-	-	-	65	-	-	15
Onychocorycaeus catus	-	-	15	-	-	-	-	-	-	-
Oncaea clevei	-	-	5	-	-	-	-	46	-	-
Oncaea media	-	-	20	-	-	-	-	-	-	-
Oncaea mediterranea	-	8	20	-	-	-	-	-	-	-
Oncaea venusta	-	-	10	-	-	-	-	-	-	-
Copilia mirabilis	-	-	-	7	-	-	-	-	-	-

Fig. 3. Taxonomic composition and abundance (individuals/m^3) of copepods across 10 sampling stations in the northern East China Sea based on morphological analysis. The taxonomic hierarchy is displayed on the left. Each row represents a detected species, with blank cells indicating species absence.

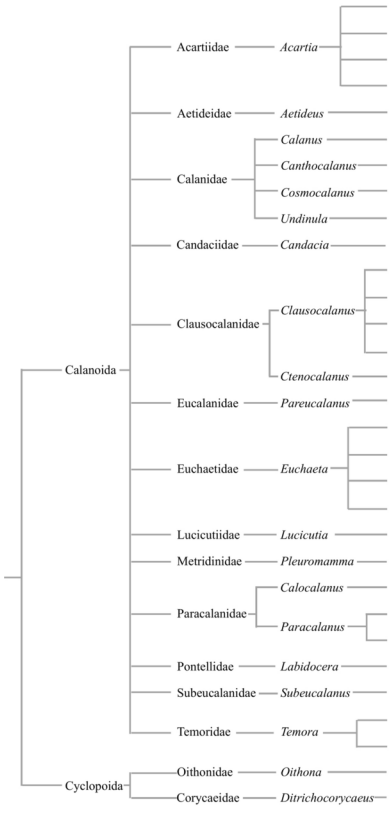
individuals to established taxonomic knowledge, allowing for immediate ecological interpretations based on known species traits and distributions. For example, our ability to distinguish between coastal and oceanic warm-water species at station EC06 provided valuable insights into the transitional nature of this sampling location. However, morphological identification has several inherent challenges. It is labor-intensive, requires substantial taxonomic expertise, and is particularly challenging for small copepods (typically 1–5 mm), where the examination of minute morphological features is required[12–14,22]. These difficulties are further exacerbated when dealing with immature stages or damaged specimens, which often lack key diagnostic features.

Another critical limitation of morphological identification is the challenge posed by cryptic speciation in copepods[8,23–26]. Cryptic species, which are morphologically similar but genetically distinct, may be overlooked, potentially leading to underestimations of biodiversity and misinterpretation of ecological patterns[27–30]. This was particularly evident in our study for closely related species pairs, such as *Acartia hudsonica* and *A. omorii*, where misidentification could result in incorrect ecological interpretations due to their differing habitat preferences.

DNA metabarcoding offered a complementary perspective, identifying 31 species from 20 genera. This molecular approach showed several advantages over traditional morphological methods. Firstly, it facilitated high-throughput analysis of bulk samples, enabling the rapid assessment of biodiversity across multiple stations. Such efficiency is particularly valuable for large-scale ecological surveys where time and resources may constrain detailed morphological analysis. Secondly, DNA metabarcoding demonstrated potential in detecting cryptic diversity that might be overlooked through morphology alone. For example, our metabarcoding analysis uniquely identified species such as *Aetideus acutus*, *Canthocalanus pauper*, and *Clausocalanus arcuicornis*, which were not detected in the morphological examination. This underscores the method's sensitivity in capturing genetic diversity that may not be evident through morphological traits.

DNA metabarcoding demonstrated robustness in identifying species across various life stages and preservation conditions. Unlike morphological identification, which often depends on adult specimens for accurate species determination, genetic methods can potentially identify individuals regardless of their developmental stage or physical condition. This advantage was evident in the stronger correlation between morphological and metabarcoding data at the genus level compared to the species level, suggesting that metabarcoding may be more reliable in classifying immature stages.

However, DNA metabarcoding also comes with its own challenges. The selection of genetic markers is crucial, as different markers offer varying levels of taxonomic resolution and coverage. In our study, we employed the



	EC02	EC04	EC06	EC08	EC10	EC12	EC14	EC16	EC18	EC20
Acartia (Acanthacartia) bifilosa	35,960	-	-	-	-	-	-	-	-	-
Acartia (Acartiura) hudsonica	-	16,450	81	2,467	-	-	-	-	-	-
Acartia (Acartia) negligens	-	-	-	80,819	-	-	-	-	-	61,680
Acartia (Odontacartia) pacifica	121,066	-	-	954	24	67,324	249,055	342,303	81,190	446
Aetideus acutus	-	-	1,159	-	-	-	-	-	-	-
Calanus sinicus	48,479	9,704	1,204	1,778	26,776	4,550	4,244	48,754	2,778	46
Canthocalanus pauper	-	-	2,142	1,633	-	-	-	218	-	50,393
Cosmocalanus darwinii	-	-	163,024	81,426	-	-	-	-	-	548,921
Undinula vulgaris	-	-	153	102,238	88	-	-	-	-	824,529
Candacia truncata	-	-	265	-	-	-	-	-	-	-
Cluasocalanus arcuicornis	-	-	193	-	-	-	-	-	-	-
Clausocalanus furcatus	-	-	646	78	-	-	-	-	-	6,261
Clausocalanus jobei	-	-	20,516	-	-	-	-	-	-	-
Clausocalanus parapergens	-	-	45,394	-	4,919	-	-	-	-	-
Ctenocalanus vanus	-	12,801	159,603	22,730	41,181	-	-	-	-	57,498
Pareucalanus attenuatus	-	-	-	-	-	-	-	-	-	5,713
Euchaeta concinna	-	-	-	-	128	-	-	-	-	480
Euchaeta indica	-	-	25,348	-	-	-	-	-	-	-
Euchaeta plana	75	23,343	9,737	28,823	134,571	-	-	-	-	25,351
Euchaeta rimana	15,268	104	19,273	46,640	15,772	-	-	-	-	51,729
Lucicutia flavicornis	-	-	197	-	-	-	-	-	-	-
Pleuromamma abdominalis	5,944	-	-	47	-	24,249	2,488,298	1,436,717	-	-
Calocalanus plumulosus	-	-	63,226	-	-	-	-	-	-	-
Paracalanus aculeatus	-	-	-	-	-	-	67,231	557,047	-	-
Paracalanus parvus s. l.	947,386	2,383,129	312,982	462,578	1,821,851	252,233	1,005,942	375,144	1,096,860	1,409,116
Labidocera acuta	-	-	-	1,412,007	-	-	69,819	87,306	-	29,521
Subeucalanus pileatus	-	-	-	7,756	-	-	-	-	-	2,365
Temora discaudata	-	-	131,669	49,505	175	16,857	919,701	1,385,560	741	1,506,370
Temora turbinata	-	-	-	-	-	14	2,056	1,123	-	-
Oithona similis	-	536,347	100,119	46,324	548,221	-	552	-	108,489	52,061
Ditrichocorycaeus affinis	34,511	19,425	2,142	32	763	2,146	9,858	4,758	5,457	1,429

Fig. 4. Taxonomic composition and relative abundance of copepods based on DNA metabarcoding across 10 sampling stations in the northern East China Sea. Values represent normalized sequence reads ($\times 10^6/m^3$). The taxonomic hierarchy is shown on the left. Each row indicates a detected species, with blank cells indicating species absence.

COI gene, which provides high species-level resolution but can be susceptible to primer bias[16,17]. This bias was evident in the underrepresentation of Cyclopoida in our metabarcoding results compared to morphological identification, suggesting potential limitations in primer binding or amplification efficiency for this order.

Another significant challenge is the reliance on comprehensive and accurate reference databases. Although databases such as NCBI offer extensive genetic information, they can include misidentifications or incomplete data, particularly for lesser-known marine taxa[31]. In our study, this limitation likely contributed to some discrepancies between morphological and metabarcoding results, especially for genera with complex taxonomy or limited genetic representation. Furthermore, the quantitative interpretation of metabarcoding data remains challenging. While we observed a moderate positive correlation between morphological abundance and sequence read counts, this relationship is influenced by various factors, including interspecific variations in DNA copy number, PCR biases, and differences in body size and biomass[32]. This complexity highlights the need for caution when inferring absolute abundances solely from metabarcoding data and emphasizes the need for systematic investigation of these methodological challenges.

To address these database-related challenges and validate our species identifications, we compared our results with a previous field study conducted in the same region[33]. Our analysis revealed high concordance in species composition between studies, with 94% (32 of 34 species) for morphological identification and 87% (27 of 31 species) for metabarcoding. This result provides strong empirical support for the reliability of our taxonomic assignments. Four species (*A. bifilosa*, *A. hudsonica*, *C. jobei*, and *S. pileatus*) were detected exclusively through our metabarcoding approach, suggesting previously unrecognized diversity in regional surveys. The high abundance of these species at specific stations, *A. bifilosa* at EC02 ($35{,}960 \times 10^6/m^3$) and *C. jobei* at EC06 ($20{,}516 \times 10^6/m^3$), indicates established populations rather than transient occurrences. These findings demonstrate metabarcoding's effectiveness in detecting cryptic species within taxonomically challenging genera such as *Acartia* and *Clausocalanus*. Conversely, our morphological analysis uniquely identified *N. robustior* and *A. gibber*, which were absent in both our metabarcoding results and previous regional survey. Their absence in metabarcoding can be attributed to previously discussed methodological limitations such as primer bias and reference database coverage. However, their exclusive detection at offshore stations (*N. robustior* at EC06 and *A. gibber* at EC20) in low abundances (15–20 individuals/m^3) requires further investigation, particularly given their absence in the previous regional study. These offshore stations are strongly influenced by the TWC, suggesting that species occurrences might be influenced by seasonal variations in oceanic circulation patterns, differences

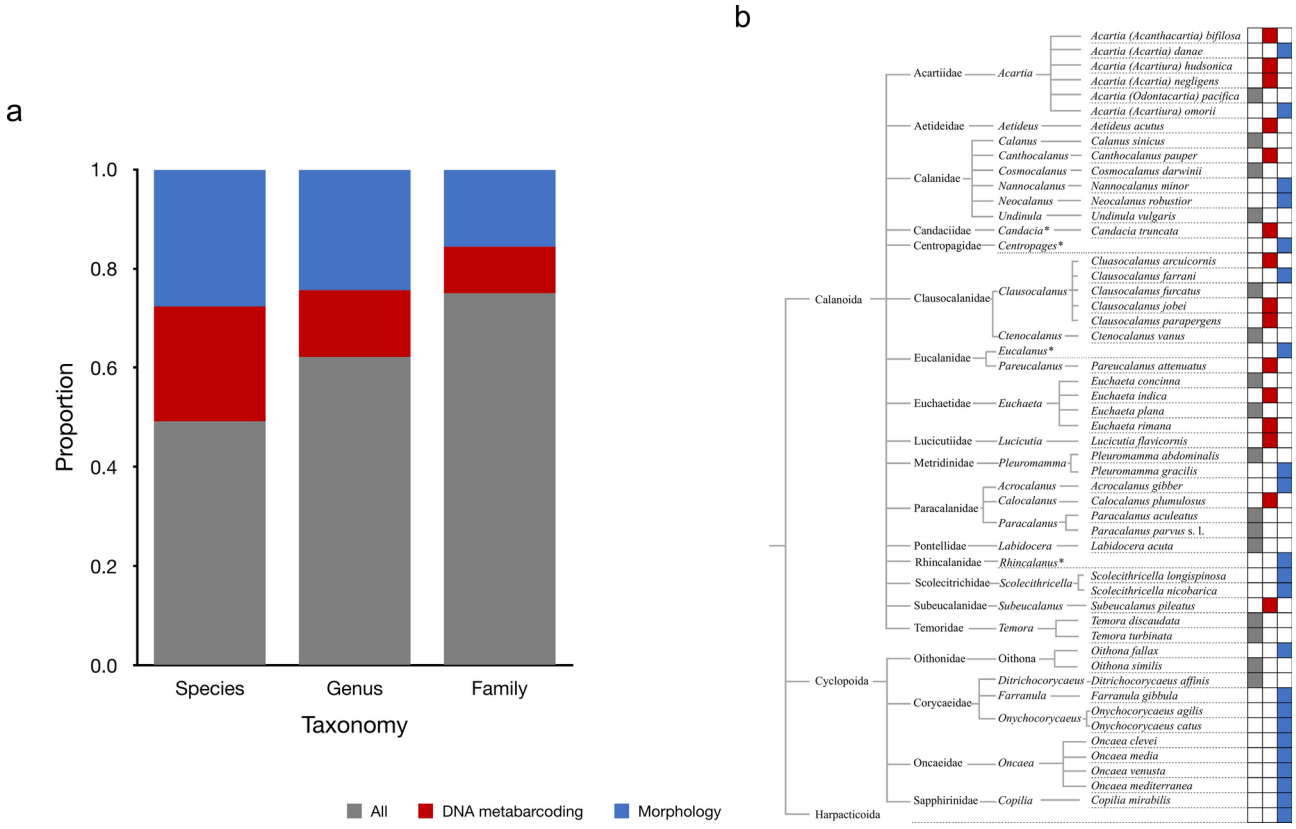


Fig. 5. Comparison of copepod identification by morphological and DNA metabarcoding approaches. (**a**) Proportion of taxa detected by each method across different taxonomic levels. (**b**) Presence-absence matrix of copepod taxa. *Genera identified only at the genus level, without species-level classification. Gray represents shared detections; red indicates taxa detected only by DNA metabarcoding; blue indicates taxa detected by morphology only.

in sampling periods relative to tidal cycles, or interannual changes in water mass distribution. Such temporal and spatial variability in environmental conditions, particularly in dynamic offshore regions where different water masses converge, could explain the discrepancy between our findings and those of the previous survey.

Despite these challenges, our systematic comparison of the two methods provided valuable insights into marine biodiversity assessment. The integration of morphological expertise with metabarcoding data revealed previously unrecognized patterns in copepod communities. Notably, we found that the concordance between methods varied systematically across taxonomic groups, with differing patterns observed between Calanoida and Cyclopoida. This order-specific variation in detection efficiency, not previously reported, provided important insights for optimizing future sampling strategies. Our detailed evaluation of primer efficiency across different copepod families further illuminated potential sources of taxonomic bias in current metabarcoding protocols. These methodological insights, combined with environmental analyses, revealed the specific limitations of each method. However, our findings also demonstrated that integrating both approaches maximized their complementary strengths, providing a more complete picture of marine biodiversity.

The integration of morphological and DNA metabarcoding approaches, informed by our methodological advances, provided a more comprehensive understanding of copepod diversity and ecology in the nECS than either method could alone. While both methods captured similar broad-scale patterns, their differences in fine-scale details offered complementary insights into copepod community structure and environmental relationships. Correlation analyses between morphology-based individual counts and DNA metabarcoding sequence read numbers provided valuable insights into the quantitative aspects of these two methods. The moderate positive correlation observed suggests that DNA metabarcoding data can, to some extent, reflect the abundance patterns detected through traditional morphological methods. Additionally, both hierarchical clustering and nMDS analyses revealed consistent patterns in copepod community structure across the two methods. The distinct grouping of stations into coastal (Group A) and offshore (Group B) clusters reflects the influence of hydrographic conditions on copepod community structure, particularly the effect of the Changjiang Diluted Water on coastal stations. The case of station EC04, classified differently by the two methods, underscores the value of employing multiple methodologies to capture the full complexity of marine ecosystems. This discrepancy may indicate a transitional zone between coastal and offshore environments, where the two methods might be sensitive to different aspects of community composition.

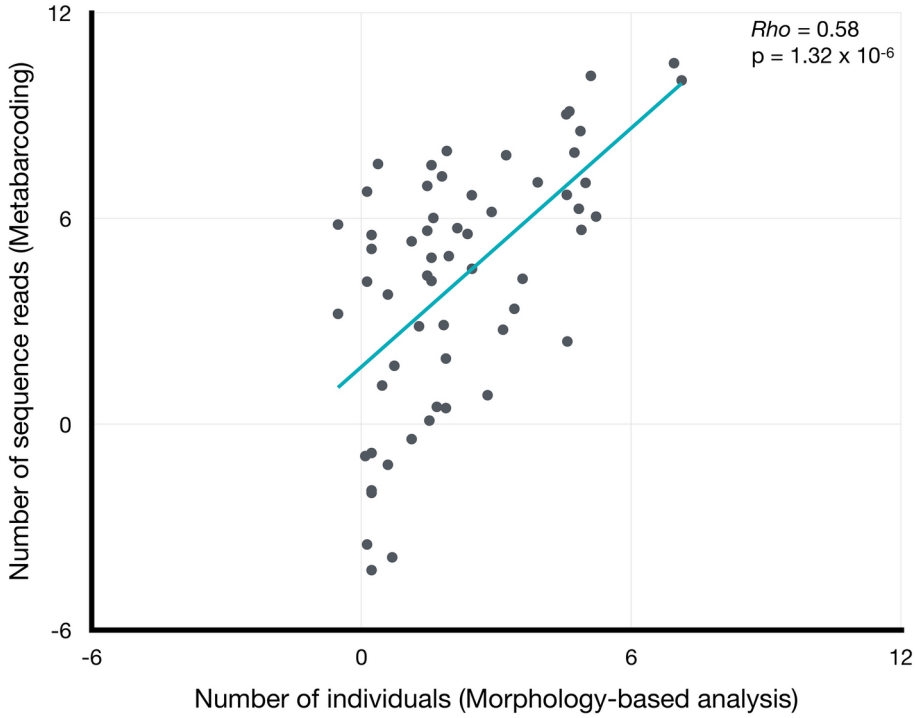


Fig. 6. Correlation between morphology-based individual counts and DNA metabarcoding sequence reads for copepod species identified by both methods. Data are centered log-ratio (clr) transformed. Spearman's rank correlation coefficient (Rho) and p-value are indicated. Each point represents a species at a sampling station.

RDA results further highlighted the complementarity of morphological and DNA metabarcoding approaches in elucidating relationships between copepod species distribution and environmental factors. Both methods consistently identified strong positive correlations between species such as *Cosmocalanus darwinii*, *Undinula vulgaris*, and *Clausocalanus furcatus* and salinity, indicating their preference for offshore, more saline waters. Conversely, species like *Acartia pacifica* and *Paracalanus aculeatus* favored coastal environments characterized by lower salinity, higher temperatures, and increased phytoplankton densities. Interestingly, some species exhibited differing environmental associations between the two methods. For instance, *Pleuromamma abdominalis* showed contrasting relationships with environmental variables in morphological and metabarcoding analyses. These discrepancies may arise from methodological differences, such as the ability of DNA metabarcoding to detect cryptic diversity or early life stages that might have different environmental preferences. Alternatively, these differences could reflect the complex life histories of these species, with distinct life stages occupying varying ecological niches.

Our study demonstrates that the integration of morphological and DNA metabarcoding approaches provides a powerful tool for the comprehensive assessment of copepod diversity and ecology. This integrated method not only enhances our understanding of marine ecosystem complexity but also informs future biodiversity monitoring strategies. Future research should continue to employ both approaches to fully elucidate the intricate relationships between copepod communities and their environments in marine ecosystems. Efforts to improve reference databases for DNA metabarcoding and to develop genetic markers that can effectively capture the full range of copepod diversity will be crucial. Additionally, integrating these methods with advanced statistical techniques and ecological modeling could further enhance our understanding of marine ecosystem dynamics and improve our ability to predict and manage changes in these critical environments. This comprehensive approach is particularly valuable in the context of global environmental changes, as it offers a more nuanced and complete picture of marine biodiversity and ecosystem functioning.

Materials and methods

Zooplankton sampling and preservation

Zooplankton samples were collected from the northern East China Sea from August 5 to 8, 2019. Sampling was conducted at ten stations (Fig. 1), covering depths from the surface to approximately 13 m. At each station, we performed vertical net hauls using a conical net with a 60 cm mouth diameter and a 200 μm mesh size. The net tows were hauled at a speed of 1 m/s to ensure a consistent sampling effort across all stations. To quantify the volume of water filtered, a calibrated flowmeter (Hydro-Bios Co., 438,115, Germany) was mounted at the net mouth. Immediately after collection, the entire zooplankton samples were preserved in 99.5% ethanol to maintain specimen integrity for subsequent analyses. This preservation method was selected to facilitate both taxonomic identification and potential molecular studies.

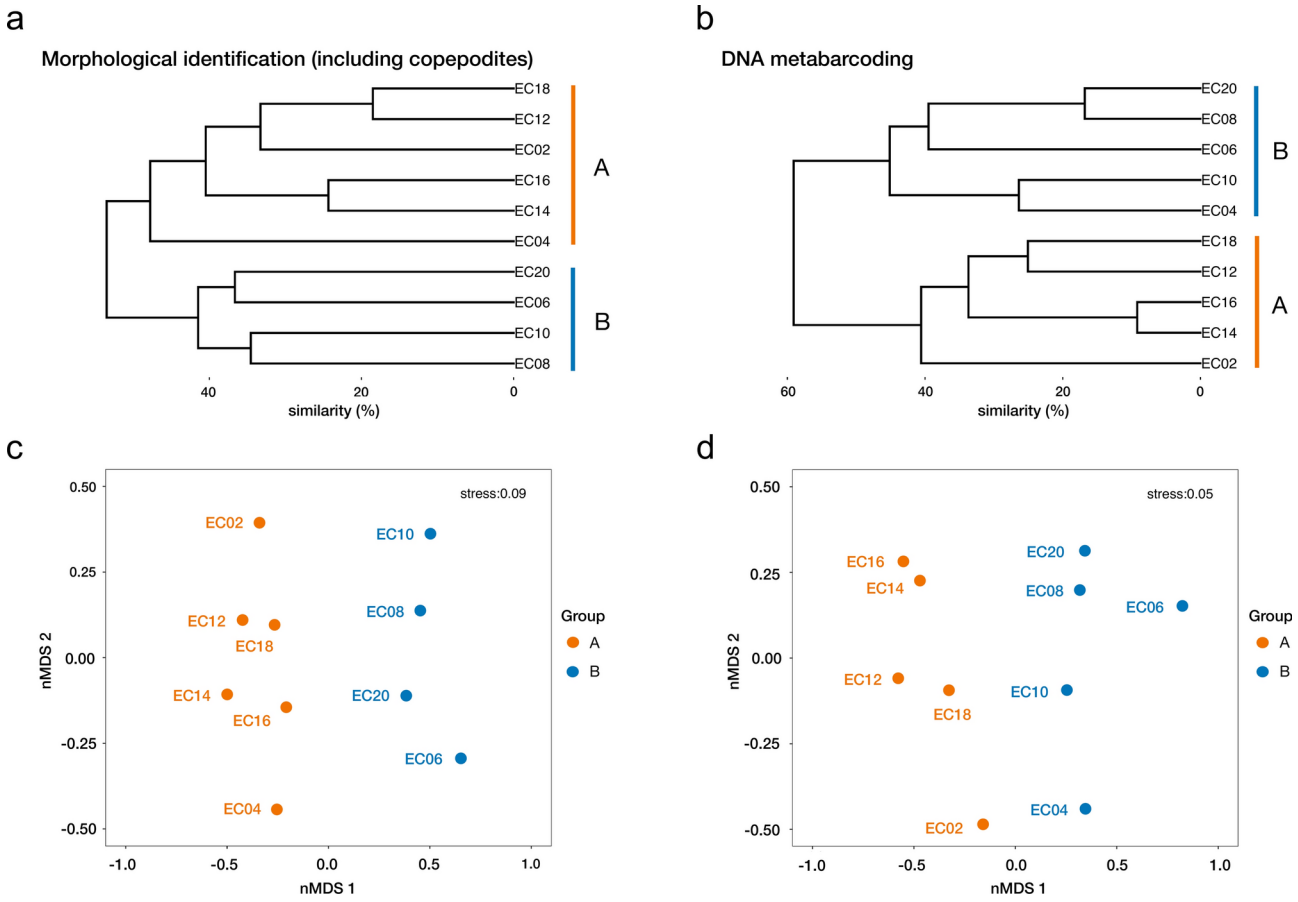


Fig. 7. Comparison of copepod community structure based on morphological identification and DNA metabarcoding. Hierarchical clustering dendrograms for (**a**) morphological and (**b**) DNA metabarcoding data. Non-metric multidimensional scaling (nMDS) plots for (**c**) morphological and (**d**) DNA metabarcoding data. Group A and B represent coastal and offshore stations, respectively. Stress values are provided for the nMDS plots.

Environmental variables measurement

Environmental factors associated with zooplankton abundance, diversity, and richness were measured at each sampling station. These factors included: (1) water temperature, salinity, and turbidity; (2) chlorophyll-a (Chl-a) concentrations; and (3) densities of diatoms, dinoflagellates, and ciliates. Vertical profiles of water temperature, salinity, and turbidity were obtained from the bottom (approximately 10 m deep) to the surface using a conductivity-temperature-depth (CTD) sensor (SBE 19plus V2, Sea-Bird Electronics, USA) and LISST-AOBS (Sequoia Scientific, Inc., USA). Concurrent with these measurements, 2 L of seawater was collected at each depth using a CTD-rosette (SBE 32, Sea-Bird Electronics, USA).

For Chl-a analysis, 1 L of the collected seawater was immediately frozen at −80 °C. Later, these samples were thawed at room temperature (approximately 20 °C), filtered through a 47 mm Whatman GF/F filter, and preserved in 90% acetone. Chl-a concentrations were measured using a UV spectrophotometer (UV-1800, Shimadzu, Japan) following the method described by Holm-Hansen et al.[34].

The remaining 1 L of seawater was fixed with 2% Lugol's solution for phytoplankton and ciliate density analysis. These samples were allowed to settle in a light-tight chamber for 48 h, after which the supernatant was carefully removed, leaving 10 mL of concentrated sample. Phytoplankton and ciliate densities were quantified using a Sedgewick-Rafter chamber (5608-C, Rigosha, Japan) under an optical microscope (Eclipse 80i, Nikon, Tokyo, Japan). Identification and enumeration were conducted following established taxonomic references[35–37].

Morphological identification and quantification of copepods

Zooplankton samples were processed for species identification and abundance estimation using standard morphological techniques. Each sample was initially divided into two equal subsamples using a Folsom plankton splitter[38]. Subsampling continued until a minimum of 200 copepod individuals were obtained, ensuring statistical reliability for abundance calculations. Copepod specimens from one subsample were identified to the species level using a combination of microscopy techniques. Initial examinations were conducted under a dissecting microscope (SMZ 645, Nikon, Japan), followed by detailed analyses of characteristic appendages using an upright microscope (Eclipse E200, Nikon, Japan). Species identification was primarily based on the

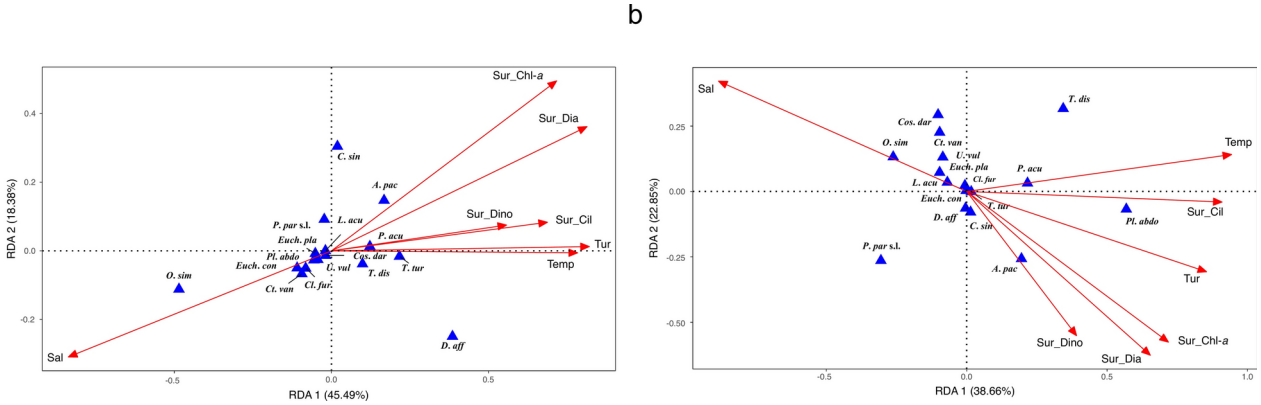


Fig. 8. Redundancy analysis (RDA) of copepod species distribution in relation to environmental factors. (**a**) Morphological identification. (**b**) DNA metabarcoding. Blue triangles represent copepod species (see abbreviations below). Red arrows indicate environmental variables. Percentages on axes indicate the variance explained by each RDA component. Abbreviations for environmental variables: Temp: Temperature (Average), Sal: Salinity (Average), Tur: Turbidity, Sur_Chl-a: Surface Chlorophyll-a, Sur_Dia: Surface Diatoms, Sur_Dino: Surface Dinoflagellates, Sur_Cil: Surface Ciliates. Abbreviations for species: A. pac: *Acartia (Odontacartia) pacifica*, C. sin: *Calanus sinicus*, Cos. dar: *Cosmocalanus darwinii*, U. vul: *Undinula vulgaris*, Cl. fur: *Clausocalanus furcatus*, Ct. van: *Ctenocalanus vanus*, Euch. con: *Euchaeta concinna*, Euch. pla: *Euchaeta plana*, Pl. abdo: *Pleuromamma abdominalis*, P. acu: *Paracalanus aculeatus*, P. par s.l.: *Paracalanus parvus* s.l., L. acu: *Labidocera acuta*, T. dis: *Temora discaudata*, T. tur: *Temora turbinata*, O. sim: *Oithona similis*, D. aff: *Ditrichocorycaeus affinis*.

illustrated guide by Chihara and Murano[36] and supplemented by our expertise in copepod taxonomy[39,40]. Key morphological features examined for accurate species identification included the segmentation and setation of antennules, the structure of mouthparts, the morphology and armature of legs, and the segmentation and ornamentation of the urosome. This comprehensive examination enabled precise differentiation between species with similar morphologies. After identification and enumeration, the abundance data were extrapolated to the entire sample and converted to individuals per cubic meter (individuals/m^3) using the volume of water filtered during sampling.

DNA extraction and metabarcoding sequencing

Genomic DNA (gDNA) was extracted from the remaining zooplankton subsample following morphological analysis. The subsample was initially centrifuged at 4000 rpm for 10 min, repeated until a clear pellet formed. The pellet was then resuspended in 50 × PBS buffer (0.5 mL per 0.5 g of pellet). Homogenization was performed on ice using an IKA-T10 homogenizer, with two 15-s intervals at 18,750 rpm followed by a 15-s interval at 22,500 rpm. Approximately 2 mL of the homogenized mixture was centrifuged at 4000×*g* for 10 min. The resulting cell pellets were resuspended in 10 mL of 2% CTAB lysis buffer, and DNA was extracted using the phenol–chloroform method. The extracted DNA was washed with 70% ethanol, precipitated, and dissolved in TE buffer (10 mM Tris–HCl, 1 mM EDTA, pH 8.0).

For copepod species identification, we targeted the mitochondrial cytochrome c oxidase subunit 1 (COI) gene, which has been widely validated as an effective marker for copepod diversity assessment due to its high resolution at the species level and extensive reference database availability. The variable region of the COI gene was amplified using the forward universal primer mICOIintF (5′-GGWACWGGWTGAACWGTWTAYCCYCC-3′) and the reverse universal primer HCO2198 (5′-TAAACTTCAGGGTGACCAAAAAATCA-3′)[41]. PCR amplification was performed in two rounds. In the first round, the reaction mixture contained 2 ng of gDNA, 5 × reaction buffer, 1 mM of dNTP mix, 500 mM of each forward and reverse primer, and Herculase II fusion DNA polymerase (Agilent Technologies, Santa Clara, CA, USA). Cycling conditions were as follows: initial denaturation at 95 °C for 3 min, followed by 25 cycles of 95 °C for 30 s, 40 °C for 30 s, and 72 °C for 30 s, with a final extension at 72 °C for 5 min. The first-round PCR products were purified using AMPure XP magnetic beads (Beckman-Coulter, Indianapolis, IN, USA). In the second round of PCR, 2 µL of the purified first-round product was used along with NexteraXT Indexed Primers for final library construction. The cycling conditions consisted of an initial denaturation step at 95 °C for 3 min, followed by 10 cycles of 95 °C for 30 s, 40 °C for 30 s, and 72 °C for 30 s, with a final extension at 72 °C for 5 min.

The final PCR products were purified, quantified using Quant-iT™ PicoGreen (Promega, Madison, WI, USA), and assessed using TapeStation D1000 ScreenTape (Agilent Technologies, Waldbronn, Germany). Paired-end sequencing (2 × 301 bp) was performed on the Illumina MiSeq platform (Illumina, San Diego, USA) by Macrogen Inc. (Seoul, Korea). Raw sequencing data were deposited in the NCBI SRA database under accession numbers SRP494366.

Metabarcoding data analysis

Raw sequence reads were preprocessed using Trimmomatic (version 0.39)[42] to remove adapter sequences and low-quality bases (Phred score < 30). Forward and reverse paired-end reads were aligned and joined using FLASH (Fast Length Adjustment of SHort reads) (v1.2.10)[43] with a minimum overlap of 10 bp. Merged reads outside the range of 300–500 nucleotides were discarded, as the expected amplicon size for the mICOIintF and HCO2198 primer pair is 313 bp[41].

Species-level identified was achieved through a hierarchical filtering approach. Initially, homology searches were performed using BLAST (v2.8.1)[44] against the NCBI non-redundant nucleotide (NT) database, with an E-value threshold of 1e-5. BLAST results were filtered based on coverage and identity criteria, calculated as follows:

$$Coverage\,(\%) = \frac{Alignment\ length}{Query\ length}$$

$$Identity\,(\%) = 1 - \frac{Number\ of\ mismatch + Number\ of\ gap\ opening}{Query\ length}$$

Sequences were retained only if they met both coverage (97%) and sequence identity (98%) thresholds. These stringent thresholds were selected based on previous studies demonstrating reliable species-level discrimination in copepods using COI markers. Retained sequences were sorted by descending bit-score, and each merged sequence was initially annotated based on the highest bit-score match.

For taxonomic assignment and verification, NCBI taxonomic identifier codes (taxIDs) were appended to the annotated sequences using the 'nucl_gb.accession2taxid' file from the NCBI taxonomy FTP server (ftp://ftp.ncbi.nlm.nih.gov/pub/taxonomy/accession2taxid/nucl_gb.accession2taxid.gz) and in-house Python scripts. Taxonomic ranks from kingdom to species were assigned and verified using TaxonKit (v 0.6.0)[45]. Sequences assigned to non-target taxa (Archaea, Bacteria, or Viruses) were excluded from further analysis. In cases where a merged sequence had multiple highest bit-score matches but different taxIDs, the sequence was manually curated and excluded if the multiple taxIDs were discordant. To ensure reliable species-level identification and minimize potential artifacts, we defined molecular operational taxonomic unit (mOTU). An mOTU was established only when at least five distinct merged sequences shared the same taxID.

Statistical analysis

Statistical analyses were performed to compare and interpret the data obtained from morphology-based and metabarcoding-based species identification methods. All analyses were conducted using R (version 4.2.1). Abundance data from both methods were standardized to individuals or sequence reads per cubic meter (individuals/m^3). For metabarcoding data, sequence reads were normalized to the total number of aligned reads in the NT library and multiplied by 10^6 to facilitate comparison with the morphological data. To account for the compositional constraint and zero-inflation typical of marine eDNA metabarcoding datasets, a centered-log ratio (clr) transformation was applied prior to correlation analysis between individual counts and sequence reads.

Community composition differences between sampling sites were visualized using non-metric multidimensional scaling (nMDS) based on Bray–Curtis distances. This analysis was performed using the 'vegan' package[46] and visualized with 'ggplot2.' The stress values, which indicate the goodness of fit between the nMDS and the original dissimilarity matrix, were 0.08 for morphology-based identification and 0.05 for metabarcoding-based identification (values range from 0 to 1, with lower values indicating a better fit). Due to negative distances resulting from centered log-transformed values, data were additionally transformed using $\log_{10}(X+1)$ for certain analyses. Hierarchical clustering analysis was conducted using the Bray–Curtis dissimilarity metric and visualized as a dendrogram with the 'factoextra' package (version 1.0.3).

Redundancy analysis (RDA) was conducted to assess the influence of environmental factors on copepod community variability. The environmental variables included temperature, salinity, turbidity, surface Chl-a concentration, and the densities of surface diatoms, dinoflagellates, and ciliates. This analysis was also conducted using the 'vegan' package.

Data availability

The raw sequencing dataset has been deposited in the NCBI SRA database under BioProject accession number PRJNA1086180.

Received: 13 October 2024; Accepted: 6 February 2025
Published online: 01 March 2025

References

1. Steinberg, D. K. & Landry, M. R. Zooplankton and the ocean carbon cycle. *Ann. Rev. Mar. Sci.* **9**, 413–444. https://doi.org/10.1146/annurev-marine-010814-015924 (2017).
2. Chen, G. & Hares, M. P. Cryptic diversity and comparative phylogeography of the estuarine copepod on the US Atlantic coast. *Mol. Ecol.* **20**, 2425–2441. https://doi.org/10.1111/j.1365-294X.2011.05079.x (2011).
3. Doney, S. C. et al. Climate change impacts on marine ecosystems. *Ann. Rev. Mar. Sci.* **4**, 11–37. https://doi.org/10.1146/annurev-marine-041911-111611 (2012).
4. Eyun, S. I. Phylogenomic analysis of Copepoda (Arthropoda, Crustacea) reveals unexpected similarities with earlier proposed morphological phylogenies. *BMC Evol. Biol.* https://doi.org/10.1186/s12862-017-0883-5 (2017).

5. Eyun, S. I., Lee, Y. H., Suh, H. L., Kim, S. & Soh, H. Y. Genetic identification and molecular phylogeny of species (Calanoida, Pseudodiaptomidae) in Korean waters. *Zool. Sci.* **24**, 265–271. https://doi.org/10.2108/zsj.24.265 (2007).
6. Goetze, E. Cryptic speciation on the high seas; global phylogenetics of the copepod family Eucalanidae. *Proc. R. Soc. B-Biol. Sci.* **270**, 2321–2331. https://doi.org/10.1098/rspb.2003.2505 (2003).
7. Keeling, R. F., Körtzinger, A. & Gruber, N. Ocean deoxygenation in a warming world. *Annu. Rev. Mar. Sci.* **2**, 199–229. https://doi.org/10.1146/annurev.marine.010908.163855 (2010).
8. Lee, C. E. & Frost, B. W. Morphological stasis in the species complex (Copepoda: Temoridae). *Hydrobiologia* **480**, 111–128. https://doi.org/10.1023/A:1021293203512 (2002).
9. Pepin, P., Colbourne, E. & Maillet, G. Seasonal patterns in zooplankton community structure on the Newfoundland and Labrador Shelf. *Prog. Oceanogr.* **91**, 273–285. https://doi.org/10.1016/j.pocean.2011.01.003 (2011).
10. Hitchcock, J. N., Mitrovic, S. M., Hadwen, W. L., Growns, I. O. & Rohlfs, A. M. Zooplankton responses to freshwater inflows and organic-matter pulses in a wave-dominated estuary. *Mar. Freshw. Res.* **67**, 1374–1386. https://doi.org/10.1071/Mf15297 (2016).
11. Furuya, K., Hayashi, M., Yabushita, Y. & Ishikawa, A. Phytoplankton dynamics in the East China Sea in spring and summer as revealed by HPLC-derived pigment signatures. *Deep-Sea Res. Part II* **50**, 367–387. https://doi.org/10.1016/S0967-0645(02)00460-5 (2003).
12. Harvey, J. B. J., Johnson, S. B., Fisher, J. L., Peterson, W. T. & Vrijenhoek, R. C. Comparison of morphological and next generation DNA sequencing methods for assessing zooplankton assemblages. *J. Exp. Mar. Biol. Ecol.* **487**, 113–126. https://doi.org/10.1016/j.jembe.2016.12.002 (2017).
13. Hirai, J., Katakura, S., Kasai, H. & Nagai, S. Cryptic zooplankton diversity revealed by a metagenetic approach to monitoring metazoan communities in the coastal waters of the Okhotsk Sea, Northeastern Hokkaido. *Front. Mar. Sci.* https://doi.org/10.3389/fmars.2017.00379 (2017).
14. Hirai, J., Kuriyama, M., Ichikawa, T., Hidaka, K. & Tsuda, A. A metagenetic approach for revealing community structure of marine planktonic copepods. *Mol. Ecol. Resour.* **15**, 68–80. https://doi.org/10.1111/1755-0998.12294 (2015).
15. Bucklin, A., Lindeque, P. K., Rodriguez-Ezpeleta, N., Albaina, A. & Lehtiniemi, M. Metabarcoding of marine zooplankton: Prospects, progress and pitfalls. *J. Plankton Res.* **38**, 393–400. https://doi.org/10.1093/plankt/fbw023 (2016).
16. Clarke, L. J., Beard, J. M., Swadling, K. M. & Deagle, B. E. Effect of marker choice and thermal cycling protocol on zooplankton DNA metabarcoding studies. *Ecol. Evol.* **7**, 873–883. https://doi.org/10.1002/ece3.2667 (2017).
17. Deagle, B. E., Jarman, S. N., Coissac, E., Pompanon, F. & Taberlet, P. DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. *Biol. Lett.* https://doi.org/10.1098/rsbl.2014.0562 (2014).
18. Lindeque, P. K., Parry, H. E., Harmer, R. A., Somerfield, P. J. & Atkinson, A. Next generation sequencing reveals the hidden diversity of zooplankton assemblages. *Plos One* **8**, e81327. https://doi.org/10.1371/journal.pone.0081327 (2013).
19. Chain, F. J. J., Brown, E. A., MacIsaac, H. J. & Cristescu, M. E. Metabarcoding reveals strong spatial structure and temporal turnover of zooplankton communities among marine and freshwater ports. *Divers. Distrib.* **22**, 493–504. https://doi.org/10.1111/ddi.12427 (2016).
20. Ershova, E. A., Wangensteen, O. S., Descoteaux, R., Barth-Jensen, C. & Præbel, K. Metabarcoding as a quantitative tool for estimating biodiversity and relative biomass of marine zooplankton. *ICES J. Mar. Sci.* **78**, 3342–3355. https://doi.org/10.1093/icesjms/fsab171 (2021).
21. Stefanni, S. et al. Multi-marker metabarcoding approach to study mesozooplankton at basin scale. *Sci. Rep.* **8**, 12085. https://doi.org/10.1038/s41598-018-30157-7 (2018).
22. Serrana, J. M., Miyake, Y., Gamboa, M. & Watanabe, K. Comparison of DNA metabarcoding and morphological identification for stream macroinvertebrate biodiversity assessment and monitoring. *Ecol. Indic.* **101**, 963–972. https://doi.org/10.1016/j.ecolind.2019.02.008 (2019).
23. Andrews, K. R., Norton, E. L., Fernandez-Silva, I., Portner, E. & Goetze, E. Multilocus evidence for globally distributed cryptic species and distinct populations across ocean gyres in a mesopelagic copepod. *Mol. Ecol.* **23**, 5462–5479. https://doi.org/10.1111/mec.12950 (2014).
24. Cornils, A., Wend-Heckmann, B. & Held, C. Global phylogeography of *Oithona similis* s.l. (Crustacea, Copepoda, Oithonidae)—A cosmopolitan plankton species or a complex of cryptic lineages?. *Mol. Phylogenet. Evol.* **107**, 473–485. https://doi.org/10.1016/j.ympev.2016.12.019 (2017).
25. Marrone, F., Lo Brutto, S., Hundsdoerfer, A. K. & Arculeo, M. Overlooked cryptic endemism in copepods: Systematics and natural history of the calanoid subgenus Borutzky 1991 (Copepoda, Calanoida, Diaptomidae). *Mol. Phylogenet. Evol.* **66**, 190–202. https://doi.org/10.1016/j.ympev.2012.09.016 (2013).
26. Rocha-Olivares, A., Fleeger, J. W. & Foltz, D. W. Decoupling of molecular and morphological evolution in deep lineages of a meiobenthic harpacticoid copepod. *Mol. Biol. Evol.* **18**, 1088–1102. https://doi.org/10.1093/oxfordjournals.molbev.a003880 (2001).
27. Aarbakke, O. N. S., Bucklin, A., Halsband, C. & Norrbin, F. Comparative phylogeography and demographic history of five sibling species of (Copepoda: Calanoida) in the North Atlantic Ocean. *J Exp Mar Biol Ecol* **461**, 479–488. https://doi.org/10.1016/j.jembe.2014.10.006 (2014).
28. Cornils, A. & Held, C. Evidence of cryptic and pseudocryptic speciation in the species complex (Crustacea, Copepoda, Calanoida). *Front. Zool.* https://doi.org/10.1186/1742-9994-11-19 (2014).
29. Halbert, K. M. K., Goetze, E. & Carlon, D. B. High cryptic diversity across the global range of the migratory planktonic copepods *Pleuromamma piseki* and *P gracilis*. *Plos One* https://doi.org/10.1371/journal.pone.0077011 (2013).
30. Laakmann, S. et al. Comparison of molecular species identification for North Sea calanoid copepods (Crustacea) using proteome fingerprints and DNA sequences. *Mol. Ecol. Resour.* **13**, 862–876. https://doi.org/10.1111/1755-0998.12139 (2013).
31. Jin, S., Kim, K. Y., Kim, M. S. & Park, C. An assessment of the taxonomic reliability of DNA barcode sequences in publicly available databases. *Algae-Seoul* **35**, 293–301. https://doi.org/10.4490/algae.2020.35.9.4 (2020).
32. Ershova, E. A., Wangensteen, O. S., Descoteaux, R., Barth-Jensen, C. & Praebel, K. Metabarcoding as a quantitative tool for estimating biodiversity and relative biomass of marine zooplankton. *ICES J. Mar. Sci.* **78**, 3342–3355. https://doi.org/10.1093/icesjms/fsab171 (2021).
33. Shin, S. S. et al. Spatiotemporal distribution characteristics of copepods in the water masses of the Northeastern East China Sea. *J. Mar. Sci. Eng.* **10**, 754 (2022).
34. Holm-Hansen, O., Lorenzen, C. J., Holmes, R. W. & Strickland, J. D. H. Fluorometric determination of chlorophyll. *ICES J. Mar. Sci.* **30**, 3–15. https://doi.org/10.1093/icesjms/30.1.3 (1965).
35. Cupp, E. E. *Marine plankton diatoms of the west coast of North America* (University of California Press, 1943).
36. Chihara, M. *An Illustrated Guide to Marine Plankton in Japan* (Tokai University Press, 1997).
37. Hallegraeff, G. M. et al. *Algae of Australia: Phytoplankton of Temperate Coastal Waters* (ABRS, 2010).
38. McEwen, G. F., Johnson, M. W. & Folsom, T. R. A statistical analysis of the performance of the Folsom plankton sample splitter, based upon test observations. *Archiv für Meteorologie Geophysik und Bioklimatologie Serie A* **7**, 502–527. https://doi.org/10.1007/BF02277939 (1954).
39. Soh, H. Y. & Moon, S.-Y. *Invertebrate Fauna of Korea, Volume 21, Number 34, Arthropoda: Maxillopoda: Copepoda: Calanoida, Cyclopoida, Marine Planktonic Copepods III_Korean* (2014).
40. Soh, H. Y., Moon, S.-Y. & Wi, J. *Invertebrate Fauna of Korea, Volume 21, Number 28, Arthropoda: Maxillopoda: Copepoda: Calanoida, Cyclopoida, Marine Planktonic Copepods II* (2013).

41. Leray, M. et al. A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Front. Zool.* https://doi.org/10.1186/1742-9994-10-34 (2013).
42. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170 (2014).
43. Magoc, T. & Salzberg, S. L. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957–2963. https://doi.org/10.1093/bioinformatics/btr507 (2011).
44. Camacho, C. et al. BLAST plus: architecture and applications. *BMC Bioinform.* https://doi.org/10.1186/1471-2105-10-421 (2009).
45. Shen, W. & Ren, H. TaxonKit: A practical and efficient NCBI taxonomy toolkit. *J. Genet. Genom.* **48**, 844–850. https://doi.org/10.1016/j.jgg.2021.03.006 (2021).
46. Dixon, P. VEGAN, a package of R functions for community ecology. *J. Veg. Sci.* **14**, 927–930 (2003).

Acknowledgements

This research was supported by the Korea Institute of Marine Science & Technology Promotion (KIMST), funded by the Ministry of Oceans and Fisheries, Korea (RS-2018-KS181192), the Global-Learning & Academic research institution for Master's · PhD students, and Postdocs (LAMP) Program of the National Research Foundation of Korea (NRF) grant funded by the Ministry of Education (No. RS-2024-00442775), and Chonnam National University (Grant number: 2018-3416).

Author contributions

S.Y.K. performed the experiments and visualization. S.C. conducted the analysis of sequencing data and drafted the manuscript. C.P. designed the experiments and supervised the conceptualization and manuscript writing. H.Y.S. supervised the conceptualization. All authors reviewed and approved the final version of the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-89684-9.

Correspondence and requests for materials should be addressed to C.P. or H.Y.S.