The crossover from microscopy to genes in marine diversity: from species to assemblages in marine pelagic copepods

Silke Laakmann1,2, Leocadio Blanco-Bercial3 and Astrid Cornils2

1Helmholtz Institute for Functional Marine Biodiversity at the University of Oldenburg (HIFMB), Ammerländer Heerstrasse 231, 26129 Oldenburg, Germany
2Alfred Wegener Institute Helmholtz Center for Polar and Marine Research, Am Handelshafen 12, 27570 Bremerhaven, Germany
3Bermuda Institute of Ocean Sciences, 17 Biological Station, GE 01 St George’s, Bermuda

An accurate identification of species and communities is a prerequisite for analysing and recording biodiversity and community shifts. In the context of marine biodiversity conservation and management, this review outlines past, present and forward-looking perspectives on identifying and recording planktonic diversity by illustrating the transition from traditional species identification based on morphological diagnostic characters to full molecular genetic identification of marine assemblages. In this process, the article presents the methodological advancements by discussing progress and critical aspects of the crossover from traditional to novel and future molecular genetic identifications and it outlines the advantages of integrative approaches using the strengths of both morphological and molecular techniques to identify species and assemblages. We demonstrate this process of identifying and recording marine biodiversity on pelagic copepods as model taxon. Copepods are known for their high taxonomic and ecological diversity and comprise a huge variety of behaviours, forms and life histories, making them a highly interesting and well-studied group in terms of biodiversity and ecosystem functioning. Furthermore, their short life cycles and rapid responses to changing environments make them good indicators and core research components for ecosystem health and status in the light of environmental change.

This article is part of the theme issue ‘Integrative research perspectives on marine conservation’.

1. Introduction

(a) Biodiversity and species identification

Biodiversity describes the variations that are found within communities, which includes variability within species, between species and between ecosystems and, as such, it is key to ecosystem functioning. Understanding biodiversity and its change constitutes the basis for conservation and management of marine biodiversity in times of perceptible changes in marine systems. For the marine pelagic realm, our current understanding of patterns in metazoan planktonic biodiversity results from decades of work by oceanographers, ecologists and taxonomists.

The identification and delimitation of species are based on various criteria that evolved from different species concepts [1]. Correct species identification is a prerequisite for most biological studies and has been traditionally based on morphological diagnostic characters. Extensive knowledge of the available reference literature and taxonomic experience are essential. To identify species independently of taxonomic expertise molecular methods have been used increasingly over the past decades. These analyses allow for a new perspective on plankton diversity and have called into question assumptions on biogeographic...
patterns and evolutionary relationships owing to the presence of cryptic or pseudo-cryptic species (sibling species with inconspicuous or non-existent morphological differences) within many taxa [2–4]. These observations imply that traditional species concepts based on morphological identified taxa may have greatly underestimated species richness [5]. Recent efforts have produced an enormous wealth of novel data from high-throughput metagenomic sampling on plankton distribution and diversity [6–9], and revealed that a large fraction of the recorded plankton diversity belongs to still unknown taxonomic groups [9]. These results illustrate that we still have too little knowledge and understanding of the diversity of plankton and their relationship to abiotic and biotic factors, especially in the light of environmental change.

To outline methodological trends in identifying, analysing and recording marine biodiversity, the present review provides an overview of the transition from morphological to molecular identification methods. For this, we use planktonic copepods as model taxon as they often dominate zooplankton communities and, as such, play a crucial role in marine systems.

### (b) Copepods as model taxon

Planktonic copepods are known for their high taxonomic and ecological diversity, making them one of the most studied marine taxonomic groups [10]. Studies comprise their biodiversity [11,12], morphology [13], taxonomy [14], phylogeny [15,16], phylogeography and distribution [17,18], life cycle strategies [19], feeding behaviour [20] or adaptation to various environmental conditions [21]. Copepods display a huge variety of behaviours, forms and life histories. They often dominate zooplankton communities, revealed by both morphological and molecular assessments, and constitute an important part in marine food webs. As such, they have an important role in the energy transfer in most marine ecosystems [22]. Owing to short life cycles and rapid responses to changing environments, they are good indicators for ecosystem health and status [23]. Therefore, the identification of copepod (and zooplankton) community composition and structure (i.e. identification of dominant species and diversity) is important to understand and to monitor changes in marine systems [23–26].

### 2. Morphological species identification and biodiversity assessments

Planktonic copepods are the most abundant metazoans on Earth [27]. In consequence, planktonic ecologists who investigate the composition or diversity of copepod communities from plankton net samples often need to identify thousands of individuals in a delimited time period. Traditionally, morphological structures are the primary tool to identify copepod species. These are usually morphological characteristics of the exoskeleton and, owing to the small size of most copepod species, these characteristics are only visible microscopically. Owing to the high number of individuals, in routine identifications only a few diagnostic characters are used to identify species [28].

Since publication of the Systema Naturae by Linnaeus [29] in 1735 more than 14 000 copepod species have been described, including more than 2000 planktonic species [27,30]. During the ‘Golden Age’ of copepod taxonomy, between the late nineteenth century to the middle of the twentieth century, large volumes emerged that were dedicated to the description of copepod species from large expeditions, e.g. [31–36]. Until today, these volumes are the foundation for species identification in marine planktonic copepods, and often their drawings are just reproduced with no changes in modern treatises. Further important volumes with species descriptions have been published in the later twentieth century, e.g. [37–42]. Identification keys exist either for copepod species of certain regions (e.g. [43,44]) or for single families or genera [42,45–48]. In 2004, a first comprehensive overview of all copepod families (also non-planktonic) was published [14] providing identification keys for genera of each family with standardized drawings. As the discovery of new species is continuing, affiliations of already described species may be subject to taxonomic revisions because when new related species are discovered, this often requires a redefinition of the taxonomic characters for the whole group, with the associated outdated of the existing keys. In consequence, the standard references are often outdated for many accepted species names and the preciseness of identification is highly dependent on the taxonomic expertise of the analyst. Species descriptions within a family or genus are often incoherent as many different authors have described the species, which complicates the evaluation of whether the specimens under consideration belong to a known species or are new to science. Furthermore, species descriptions are generally based on morphological characters of adult specimens and are also gender-related, which makes it nearly impossible to identify early (nauplii) and juvenile (copepodite) life stages, both more abundant than adults. Lists of currently accepted species are compiled and updated at the World Register of Marine Species (WoRMS, [29]) and the Marine Planktonic Copepods webpage [49], with the latter also including drawings and biogeographic notes for each species.

A limitation on the species discrimination and definition based on morphological characters is the subjective nature of diagnostic characters. Unless crossing experiments are carried out, the definition of a species’ limits, identity and associated diagnostic characters are always subject to the criteria of the taxonomist. That is, whatever characters are selected as those that draw the line between species rather than mere variability or morphotypes within a species depends not only on the data available to the researcher, but also on the researcher’s subjective opinion of what a species is. Furthermore, the discrimination of sibling species is often not even possible (e.g. within the prominent genus Calanus), at least using characters that could be used on a routine basis without resourcing to complicated microscopy procedures such as the study of tegumental pores [50,51]. Thus, rare species may be overlooked or co-occurring sister species that differ only in minuscule characteristics may be merged.

The sampling method also has a great impact on the preciseness of identification. Often swimming and mouth appendages hold the features that characterize species in copepods. In net samples, these appendages are often broken and may thus only allow identification of the specimen to the genus level. Large mesh sizes (greater than 200 µm) have also led to more intensive studies on larger calanoid copepods than on the often more abundant cyclopoid copepods [52,53]. Recent geometric morphometric approaches allow the avoidance of problems arising from missing morphological characters [54], but are difficult to implement in routine identifications. For all these reasons, juveniles and non-calanoid copepods are often grouped to higher taxonomic levels [55],
and starting by the end of the past century, several molecular and morphological identification of species, challenges using morphological identification of species, To help in addressing one or several of the previous questions (a) Molecular identification of single species many different taxonomic entities [58].

Species identification is, however, hampered by the condition of the organism, and quantification (abundance, biomass) that the regions targeted by the restriction enzyme or those conserved within-species variability. These techniques are standardized identification, automation which results in underestimating species diversity and richness. To facilitate routine identification, techniques based on machine learning have been developed to semi-automatically identify and quantify the composition of plankton assemblages from images of preserved samples at a relatively coarse taxonomic level (ZooScan [56], EcoTaxa [57]). These techniques extract not only useful information on abundance, but also several metrics that allow the estimation of individual sizes.

In summary, morphological species identification of copepod samples not only enable the study of taxonomic diversity but also provide information on abundances, biomass, size class distribution and life stage composition (table 1). Next to this, the study of organisms allows the collection of data on species traits and thus functions in marine systems. Species identification is, however, hampered by the condition of the organisms and the level of taxonomic expertise. If several analysts with a different experience level identify species from sets of samples, e.g. in long-term monitoring efforts, the list of species and stages counted and identified vary with taxonomic changes and increasing expertise, resulting in many different taxonomic entities [58].

3. Molecular techniques to address diversity (a) Molecular identification of single species To help in addressing one or several of the previous questions and challenges using morphological identification of species, and starting by the end of the past century, several molecular methodologies have been developed for the study of Copepoda. The first methods used for species discrimination were based in variations of fragment length analyses. Most of the methods were developed in the 1980s in the biomedical field [59,60] and successfully applied to other fields soon after their discovery, including many applications to the species or lineages discrimination. For copepods, some of the first groups to benefit from these methods were the ecologically relevant Pseudocalanus [61,62] and the North Atlantic Calanus species complexes [62–64] and later on other species complexes at local or regional level [65,66]. Two main methods were used for these pioneering studies. Meanwhile, a species-specific PCR, based on competitive priming between species-specific primers [60], was the method used in some studies [60–62,64] and a restriction fragment length polymorphism (RFLP; [59]) was the approach developed by others [63]. All studies were developed on mitochondrial genes (16S rRNA, cytochrome c oxidase subunit I (COI)) to take advantage of the existing high between-species variability and relatively conserved within-species variability. These techniques are robust, time-efficient and of low costs (both in terms of daily expenses (consumables) and equipment required, with just a thermocycler and a gel system needed). On the other hand, to develop a reliable method, an in-depth knowledge of the genetic diversity of the studied species is needed, to ensure that the regions targeted by the restriction enzyme or those binding to the species-specific oligonucleotides are highly conserved within the species. In marine copepods, with population sizes often in the range of billions to trillions, it

| potentials | drawbacks |
|------------|-----------|
| morphological identification | # requires taxonomic expertise on diverse groups |
| # information on | # incoherent species descriptions |
| — life stage composition and size class distribution | # gender- and stage-related diagnostic species characters |
| — traits and hence ecological role/function | # identification depends on condition of the organism |
| — quantification (abundance, biomass) | # subjective nature of diagnostic characters |
| molecular genetic identification | # no identification of sibling and cryptic species and populations |
| single species | # time intensive |
| # species identification of young developmental stages and cryptic/sibling species → higher diversity | # requires prior methodological knowledge |
| # identification of populations | |
| # standardized identification, automation | |
| metabarcoding | # no information on |
| # simultaneous identification of a multitude of species | — community structure regarding size and stage distribution |
| # processing of large numbers of samples | — biomass and abundance |
| # efficient and cost-effective for analysing bulk samples | — ecological role |
| # standardized identification, automation | # depends on high-quality sequence reference database for different regions and progress in providing sequence reference entries |
| | # identification of thresholds |
| | # false positives, false negatives |
| | # primer and PCR biases (amplicon sequencing) |

Table 1. Potentials and drawbacks for the traditional morphological and molecular genetic species identification in biodiversity analyses.
would be very difficult to ensure that all individuals would show a conserved region long enough to fit the enzyme site or the oligonucleotide region [66] and coverage should include all known populations to consider private alleles. Furthermore, they are not easily expandable, and the addition of any new species would usually involve having to develop the protocol from scratch. Another problem associated with these methods is the lack of ability to detect cryptic species, since they can be mistaken for one of the existing species (if the target region(s) is/are identical) or for a negative result (if none of the regions are conserved). Despite the rise of DNA-sequence based methods soon after (which overcome some of the aforementioned problems), still some related methods have been developed in recent years. Amplicon length variability, based on insertion/deletion markers, has been used to discriminate between all North Atlantic Calanus species [67,68]. This method comprises a number of different indel regions, and is robust against the failure of one of the markers in case of priming site variability, since the remaining markers would allow species assignment. With the advantage of little time and budget investment required, these methods were used, for example, to characterize the distribution of the different Calanus species in the North Atlantic [68,69].

Similarly, the analysis of DNA fragment lengths such as multilocus microsatellite fingerprinting was also used to discriminate between sister species [70] and in combination with DNA sequencing of mitochondrial genes, to support the presence of cryptic speciation within a taxonomically complex species [71]. This method gives insights between sister species in a greater genetic resolution and allows parallel studies on gene flow and dispersal. However, the major drawback of this approach is the intensive development of microsatellite markers specific for every species, and the need to re-develop the method from scratch when adding and combining several species to avoid ascertainment bias.

More recently, the use of genome-wide single nucleotide polymorphisms (SNPs) has been opening a new door to discriminating between cryptic species clusters with high genetic diversity and sympatric mitochondrial DNA clades [72].

With the refinement of DNA sequencing, which is known nowadays as DNA barcoding, the species identification based on the sequence of a relatively short fragment of DNA [73,74] was developed in the early 2000s. This method was already used previously to differentiate between species or forms of marine copepods [2,75–77]. Although different authors made use of a number of different markers, especially the mitochondrial genes (16S, COI), with the launch of the Consortium for the Barcode of Life initiative [74], the use of the Folmer region [78] of the COI was the chosen for metazoan (copepods included) barcoding. The original objective of DNA barcoding was not to address taxonomic questions (DNA taxonomy) but just developed as an identification tool (see review by [79]). Within this scope, several initiatives were oriented to provide a database of known species, providing global or regional molecular references based on morphologically identified individuals by taxonomic experts, often flagging some potential cryptic speciation issues [12,80,81], a hidden diversity that could not be addressed by morphological methods (table 1).

But, even before the launch of the barcoding initiatives, the use of sequences was often oriented as a tool to aid in solving taxonomic problems—ideally as a complement to morphological studies [82]—to understand the cryptic diversity within Copepoda, which is indiscernible by morphological characters. Compared to fragment length-based methods, the use of DNA sequences (independently of the marker) allowed scientists to study other facets of the biology and the taxonomy of species, such as degree of relatedness between species (by distance methods or phylogenetic reconstructions), to detect, reject or support the presence of cryptic species, and even to delineate genetically isolated subpopulations within a species. Many of the previously mentioned molecular studies require detailed morphological information and the combination of both is nowadays known as integrated taxonomy. Within this approach, information on both the morphological and molecular species identification is paired at individual level, and ideally stored in open-source sequence reference libraries, such as BOLD (http://www.boldsystems.org/) or, (even better), paired with a museum collection. Such an approach has been very fruitful for characterizing cryptic diversity of open ocean copepods [2,4,83–85], a diverse range of species complexes [3,86,87] especially when they are used as important indicators of climate responses [68,88], meso- and bathypelagic hidden diversity [89], the identity of key players in upwelling ecosystems [90–92], and for dealing with the always extreme complexity of non-calanoid copepods [18,93], for which the relevance in the ocean ecosystems has been always been understudied owing to the complexity of their taxonomy [53]. Meanwhile, while molecular methods alone are useful to detect isolated evolutionary lineages [12,94], without an accompanying taxonomic and morphological study it would be very difficult to use this information further to infer the ecological relevance of such hidden diversity, especially in the context of past studies.}

(b) Molecular identification of assemblages and communities

With the advent of high-throughput sequencing techniques, the molecular genetic identification of single metazoan species moved fast forward towards the analysis of whole marine metazoan communities such as meiofauna [95] or zooplankton [6,96]. Multiple species and entire communities, can be identified simultaneously by analysing orthologous gene regions in parallel from environmental samples using next-generation sequencing platforms. This process is defined as metabarcoding. Compared to DNA barcoding on single specimens, metabarcoding is based on shorter gene fragments, and in general of a single marker, often variable regions of conserved nuclear small-subunit ribosomal RNA genes 18S rRNA (V1-2 [6,96,97]; V4 [98,99]; V9 [9,100–107] and 28S rDNA [7,8]). Owing to their conserved nature, resolving the obtained sequences to identify species is often impossible, since the same sequence for that region might be shared between genera, families or even superfamilies, depending on the marker used and the phylogenetic divergence between the different species. Mitochondrial markers allow a better taxonomic resolution and species identification compared to nuclear markers [108–112]. However, owing to the less conserved primer regions in COI it also implies primer mismatches and consequently missing amplification of a wide range of taxa (false negatives). Possible solutions to this problem are multi-marker and multi-primer approaches [111,113] as they enable the identification to different taxonomic levels and of a greater proportion of different taxa and thus biodiversity.

Successful species assignment also requires a complete and high-quality reference sequence database, ideally for the location and the season. For copepods, there are a limited
number of high-quality DNA barcodes available for the identification [12,79,114] and such a shortage may lead either to misidentification or to an underestimation of diversity owing to non-identification of the sequences obtained [100,110]. When DNA reference sequences do not exist and sequences cannot be identified, similarity/divergence thresholds are chosen to cluster sequences into taxonomic units [95,98,110]. The choice of the threshold depends on the divergence in the chosen gene fragment between species, genera and families in closely- and distantly-related organisms, and has an influence on the number or taxa detected by metabarcoding.

Despite these discussed challenges, metabarcoding provides a new and more comprehensive view on zooplankton and copepod biodiversity and assemblages by detecting hidden diversity and by offering the possibility of automation and a cost-effective analysis to be applied in time-series analysis and ocean ecosystem assessments. Metabarcoding analyses have been a major breakthrough in identifying bulk samples of zooplankton and thus copepod assemblages as they detect many different species from diverse taxa, cryptic species or developmental stages (e.g. meroplankton, nauplii, copepodites), previously hidden by morphological identification only (see §2), resulting in higher number of taxa or diversities (figure 1) [6,7,96,105,107,110,115]. Moreover, rare and non-indigenous species, among others, are more likely to be identified by a molecular genetic approach [103]. Several studies on zooplankton have also demonstrated the ability of metabarcoding to map both temporal and spatial patterns in zooplankton diversity and assemblages [8,100,106,107,110,111].

A major drawback of metabarcoding of bulk samples compared to morphological analysis is that quantitative ecological parameters that characterize communities such as the composition of life stages, abundances and biomass cannot currently be assessed. For such a quantitative analysis, the correlation between the number of sequence reads and the biomass or the abundance of the individual species is still in discussion, since biases might exist owing to, for example, the differences in gene copy numbers, or to PCR bias, in which the chosen primer may match better in some taxa and thereby leads to a better amplification of these specific taxa (primer match/mismatch). Despite these caveats, analysing mock assemblages of pelagic copepods (on family level) showed a significant correlation of the number of sequence reads to the dry weight of the taxon [7]. In real samples, sequence numbers and counts of calanoid copepods showed a significantly positive correlation [107] and a high correspondence in whole zooplankton samples [103]. Since the number of reads would only reflect relative abundances (there is no relationship between amount of DNA extracted and number of reads obtained after sequencing) some other analyses (DNA quantification by quantitative PCR, addition of internal DNA standards at DNA extraction or amplification, identification of mock communities, biomass measures, among others) would still be needed to move from read number to a biomass-like measurement.

Next to genetic material from whole zooplankton samples, multiple species can also be identified from genetic material sampled from the environment, referred to as environmental DNA (eDNA). By definition, eDNA is the DNA extracted from an environmental sample such as water, soil or air without isolating the target organism [116]. Large-scale biodiversity surveys based on eDNA and using both nuclear ribosomal and/or mitochondrial markers have shown that this methodology can provide valuable insights in the zooplankton and copepod communities, including potential invaders [117,118].

A comparative approach on zooplankton, including morphological identification and multi-marker metabarcoding of bulk samples and eDNA, revealed significant differences in taxonomic compositions. However, the dominant copepod taxa (identified to the family level) were identified in all of the three different approaches [109]. Metabarcoding of bulk samples gave a better measure of the zooplankton (and the morphological identification) itself, but eDNA metabarcoding better reflected the overall diversity of the broader marine community, which is not as accessible and easy to sample as zooplankton. To improve the detection of organisms in metabarcoding of eDNA, best-practices, ranging from field to laboratory and data processing standards, are still needed. These would include minimum reporting standards regarding study design, water collection, sample preservation, extraction process and high-throughput sequencing [119]. Furthermore, raw data (FASTQ files) and processing data pipelines should be made available for the scientific community by the storage in complementary repositories to allow full transparency and reproducibility (outlined by [120]).

4. Conclusion
Our understanding of pelagic biodiversity results from decades of morphological taxonomic work; however, this knowledge
has to be integrated into novel molecular genetic approaches. Only by adjusting and linking these new tools with the traditional methods can we maintain the acquired knowledge for future research using molecular data as the main workhorse for community ecology and taxonomy. The power of a so-called total evidence approach by identifying plankton, relying on both molecular and morphological information whenever possible (also referred to as ‘successful marriage of molecular and morphological methods’) was already outlined one decade ago [28]. At that time, the authors outlined: It has also been suggested that advances in sequencing technologies may overtake the single-gene barcode approach by enabling rapid genomics of species, even during routine sampling, making the current mitochondrial-based barcoding seem not ambitious enough [28, p. 1121]. Nowadays, in the times of high-throughput sequencing techniques, we are able to analyse whole communities based on molecular data, but it is still advisable to integrate morphological approaches to ensure identification, and especially quantification. For the simultaneous identification of multiple species in whole samples, the comparison between morphological and molecular identification by metabarcoding confirms that the molecular approach is not yet ready to completely replace traditional taxonomy by morphological analyses [110]. Hence, metabarcoding still needs the ground truthing by the direct comparison to the traditional morphological taxonomic analysis of the sample [107], especially for quantification [121].

For zooplankton, molecular genetic studies demonstrate promising diversity analyses based on bulk samples and allow the processing of large numbers of samples, which is advisable (in combination with traditional methods, table 2) for future studies on planktonic communities. However, three checkpoints should be implemented in future metabarcoding studies on plankton: protocol optimization, error minimization and a downstream analysis that considers potential and remaining biases [121]. To analyse marine life across groups, communities, taxa or environments that are not as accessible and easy to sample as zooplankton, eDNA analyses constitute a promising, non-invasive and non-destructive methodology to provide insights into marine life, especially when the organisms are rare, big, elusive, threatened, endangered, non-indigenous or cryptic. Particularly for metabarcoding of eDNA to detect macroorganisms, we need best-practices ranging from field to laboratory and data processing standards [121].

Looking back on the fast development and improvement in the field of species, assemblage and community identification, the continuous methodological advancement of sequencing technologies and bioinformatics will, if validated by traditional methods, allow a more comprehensive view on marine life to be included in marine conservation and management.

**Data accessibility.** This article has no additional data.

**Authors’ contributions.** All authors contributed to the conception and design of this review article. The different sections were drafted individually and were then critically revised by the other authors. All authors approved the version to be published.

**Competing interests.** We declare we have no competing interests.

**Funding.** HIFMB is a collaboration between the Alfred-Wegener-Institute, Helmholtz-Center for Polar and Marine Research, and the Carl-von-Ossietzky University Oldenburg, initially funded by the Ministry for Science and Culture of Lower Saxony and the Volkswagen Foundation through the ‘Niedersächsisches Vorab’ grant program (grant no. ZN3285). The work of WG 157 presented in this article results, in part, from funding provided by national committees of the Scientific Committee on Oceanic Research (SCOR) and from a grant to SCOR from the US National Science Foundation no. (OCE-1840868).

**Acknowledgements.** This review is a product of SCOR Working Group MetaZooGene (WG157) and Working Group on Integrated Morphological and Molecular Taxonomy (WG157) of the International Council for the Exploration of the Sea (ICES). We thank both groups for facilitating this research. We want to thank the editor and anonymous reviewers for their helpful comments and suggestions on the manuscript.

### Table 2. Guidelines to select approach to specific research questions.

| research question | approach |
|--------------------|----------|
| cladistics and systematics | integrative analyses |
| ecological studies on community structure | morphological identification |
| single/few sister species distribution | morphological identification, fragment analysis |
| alpha-diversity monitoring | metabarcoding |
| invasive species | metabarcoding (bulk samples and eDNA) |

### References

1. De Queiroz K. 2007 Species concepts and species delimitation. *Syst. Biol.* **56**, 879–886. (doi:10.1080/10635150701701083)

2. Goetze E. 2003 Cryptic speciation on the high seas; global phylogenetics of the copepod family Eucalanidae. *Proc. R. Soc. B* **270**, 2321–2331. (doi:10.1098/rspb.2003.2505)

3. Cornils A, Held C. 2014 Evidence of cryptic and pseudocryptic speciation in the Paracalanus parvus species complex (Crustacea, Copepoda, Calanoida). *Front. Zool.* **11**, 19. (doi:10.1186/1742-9994-11-19)

4. Bode M, Laakkam S, Kaiser P, Hagen W, Auel H, Cornils A. 2017 Unravelling diversity of deep-sea copepods using integrated morphological and molecular techniques. *J. Plankton Res.* **39**, 600–617. (doi:10.1093/plankt/fbx031)

5. Norris RD. 2000 Pelagic species diversity, biogeography, and evolution. *Paleobiology* **26** (4, Suppl.), 236–258. (doi:10.1666/0094-8373(2000)26[236:PSBAAE]2.0.CO;2)

6. Lindeque PK, Parry HE, Hammer RA, Somerfield PJ, Atkinson A. 2013 Next generation sequencing reveals the hidden diversity of zooplankton assemblages. *PLoS ONE* **8**, e81327. (doi:10.1371/journal.pone.0081327)

7. Hirai J, Kuriyama M, Ichikawa T, Hidaka K, Tsuda A. 2015 A metagenetic approach for revealing community structure of marine planktonic copepods. *Mol. Ecol. Res.* **15**, 68–80. (doi:10.1111/1755-0998.12294)

8. Hirai J, Tsuda A. 2015 Metagenetic community analysis of epipelagic planktonic copepods in the tropical and subtropical Pacific. *Mar. Ecol. Prog. Ser.* **534**, 65–78. (doi:10.3354/ meps11404)
copepods (Crustacea) for interpreting inter-species relationships. Mar. Biol. 120, 107–114.

52. Paffenhofer GA, Mazzaoci MG. 2003 Vertical distribution of subtidal epipelagic copepods. J. Plankton Res. 25, 1139–1156. (doi:10.1093/plankt/25.9.1139)

53. Gallienne CF, Robins DB. 2001 Is Oithona the most important copepod in the world’s ocean? J. Plankton Res. 23, 1421–1432. (doi:10.1093/plankt/23.12.1421)

54. Karamovic I, Lee S, Lee W. 2018 Instant taxonomy: choosing adequate characters for species delimitation and description through congruence between molecular data and quantitative shape analysis. Invert. Systems 32, 551–580. (doi:10.1071/IS17002)

55. Cornils A, Schulz J, Schmitt P, Lanuru M, Richter C, Schnack-Siel1 SB. 2010 Mesozooplankton distribution in the Spermonde Archipelago (Indonesia, Sulawesi) with special reference to the Calanoida (Copepoda). Deep-Sea Res. Part II: Top. Stud. Oceanogr. 57, 2066–2088. (doi:10.1016/j.dsr2.2010.09.011)

56. Gorsky G et al. 2010 Digital zooplankton image analysis using the ZooScan integrated system. J. Plankton Res. 32, 285–303. (doi:10.1093/plankt/fbp124)

57. Picheral M, Colin S, Iriasa J-O. 2017 EcoTaxa, a tool for the taxonomic classification of images. http://ecotaxa.obs-vlfr.fr.

58. Greve W, Reiners F, Hoarau G, Lindeque P, Castellani C, Møller EF, Blanco-Bercial L. 2004 RFLP procedure to discriminate between Clausocalanus Gresbrecht, 1888 (Copepoda, Calanoida) species in the Central Cantabrian Sea. J. Exp. Mar. Biol. Ecol. 344, 73–77.

59. Smolina I, Kollias S, Poortvliet M, Nielsen TG, Lindeque P, Castellani C, Müller EF, Blanco-Bercial L, Hoarau G. 2014 Genome- and transcriptome-assisted development of nuclear insertion/deletion markers for Calanus species (Copepoda: Calanoida) identification. Mol. Ecol. Res. 14, 1072–1079. (doi:10.1111/1755-0996.12241)

60. Chouquet M et al. 2017 Genetests redraws pelagic biogeography of Calanus. Biol. Lett. 13, 20170588. (doi:10.1098/rsbl.2017.0588)

61. Lindeque P, Harris RP, Jones MB, Smerdon GR. 2004 Multilocus evidence for globally competitive, species-specific PCR. Nucleic Acids Res. 32, 471–508. (doi:10.1093/nar/gnh049)

62. DeSalle R, Egan MG, Siddall M. 2005 The Unholy Trinity: taxonomy, species delimitation and DNA barcoding. Phil Trans. R. Soc. B 360, 1905–1916. (doi:10.1098/rstb.2005.1722)

63. Goetze E, Bradford-Grieve J. 2005 Genetic and morphological description of Eucalanus spinifer T. Scott, 1894 (Calanoida: Eucalanidae), a circumglobal sister species of the copepod E. hyalinus s.s. (Claus, 1866). Prog. Oceanogr. 65, 55–87. (doi:10.1016/j.pocean.2002.03.018)

64. Hill RS, Allen LD, Bucklin A. 2001 Multiple species-specific PCR protocol to discriminate four N. Atlantic Calanus species, with an mtCOI gene tree for ten Calanus species. Mar. Biol. 139, 279–287. (doi:10.1007/s0022700100548)

65. Kempter J, Piasecki W, Węski K, Krawczyk B. 2006 Systematic position of copepods of the genus Achetae (Crustacea: Copepoda: Stiphonostomatotida) parasitizing perch, Perca fluviatilis L., and zander, Sander lucioperca (L.). J. Fish Dis. 29, 103–111. (doi:10.1111/j.1365-2761.2006.00693.x)

66. Blanco-Berial L, Álvarez-Marqués F. 2007 RFLP procedure to discriminate between Clausocalanus Gresbrecht, 1888 (Copepoda, Calanoida) species in the Central Cantabrian Sea. J. Exp. Mar. Biol. Ecol. 344, 73–77.

67. Andrews KR, Norton EL, Fernandez-Silva I, Porter E, Goetze E. 2014 Multilocus evidence for globally distributed cryptic species and distinct populations across ocean gyres in a mesopelagic copepod. Mol. Ecol. 23, 5462–5479. (doi:10.1111/mec.12950)

68. Hirai J. 2019 Insights into reproductive isolation within the pelagic copepod Pleuronamama abnormalis with high genetic diversity using genome-wide SNP data. Mar. Biol. 167, 1. (doi:10.1007/s00227-019-13618-x)

69. Hebert PDN, Ratnasingham S, de Waard RJ. 2003 Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc. R. Soc. Lond. B (Suppl.) 270, 596–599. (doi:10.1098/rspb.2003.0035)

70. Hebert PDN, Cywinska A, Ball SL, DeWaard JR. 2003 Biological identifications through DNA barcodes. Proc. R. Soc. B 270, 313–321. (doi:10.1098/rspb.2002.2218)

71. Bucklin A, Frost WB, Kocher TD. 1992 DNA sequence variation of the mitochondrial 16S rRNA in Calanus (Copepoda; Calanoida): intraspecific and interspecific patterns. Mol. Mar. Biol. Biotech. 1, 397–407.

72. Bucklin A, Laleunesse TC. 1994 Molecular genetic variation of Calanus pacificus (Copepoda: Calanoida): preliminary evaluation of genetic structure and subspecific differentiation based on mtDNA sequences. CaCOOp Rep. 35, 45–51.

73. Bucklin A, Frost WB, Kocher TD. 1995 Molecular systematics of six Calanus and three Metridia species (Calanoida: Copepoda). Mar. Biol. 121, 655–664. (doi:10.1007/BF00349301)

74. Frolich O, Black M, Hoew L, Lutz R, Vrijenhoek R. 1994 DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Mar. Biol. Biotech. 3, 294–299.

75. Bucklin A, Steinke D, Blanco-Berial L. 2011 DNA barcoding of marine metazoa. Annu. Rev. Mar. Sci. 3, 471–508. (doi:10.1146/annurev-marine-120308-080950)

76. Bucklin A, Ortmann BD, Jennings RM, Nigro LM, Sweetman CJ, Copley NJ, Sutton T, Wiebe PH. 2010 A “Rosetta Stone” for metazoan zooplankton: DNA barcode analysis of species diversity in the Sargasso Sea (Northwest Atlantic Ocean). Deep-Sea Res. Part II 57, 2234–2247. (doi:10.1016/j.dsr2.2010.09.025)

77. Bucklin A, Hopcroft RR, Kosobokova KN, Nigro LM, Ortmann BD, Jennings RM, Sweetman CJ. 2010 DNA barcoding of Arctic Ocean holozooplankton for species identification and recognition. Deep-Sea Res. Part II 57, 40–48. (doi:10.1016/j.dsr2.2009.08.005)

78. DeSalle R, Egan MG, Siddall M. 2005 The Unholy Trinity: taxonomy, species delimitation and DNA barcoding. Phil. Trans. R. Soc. B 360, 1905–1916. (doi:10.1098/rstb.2005.1722)

79. Stupnikova AN, Molodtsova TN, Mugue NS, Neretina TV. 2013 Genetic variability of the Metridia lucas complex (Copepoda) in the Southern Ocean. J. Mar. Sys. 128, 175–184. (doi:10.1016/j.jmarsys.2013.04.016)

80. Ueda H, Bucklin A. 2006 Acanth C. cf. (Odontacartia) obtusikia, a new brackish-water copepod from Ariake Bay, Japan, with a redescription of the closely related A. pacifica from the Seto Inland Sea. Hydrobiologia 560, 77–91. (doi:10.1007/s10750-005-9513-0)

81. Evun S-I, Lee Y-H, Suh H-L, Kim S, Suh HY. 2007 Genetic identification and molecular phylogeny of Pseudodiaptomus species (Calanoida, Pseudodiaptomidae) in Korean Waters. J. Jpn. Soc. 24, 265–271. (doi:10.2108/zjsj.24.265)

82. Lindeque PK, Hay SJ, Heath MR, Ingvars'dottir A, Rasmussen J, Smerdon GR, Wanickel J. 2006 Integrating conventional microscopy and molecular analysis to analyse the abundance and distribution of four Calanus congeners in the North Atlantic. J. Plankton Res. 28, 221–238. (doi:10.1093/plankt/fbt115)
