Vol. 534: 65–78, 2015
doi: 10.3354/meps11404

MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser

Published August 27

Metagenetic community analysis of epipelagic planktonic copepods in the tropical and subtropical Pacific

Junya Hirai[1,*], Atsushi Tsuda[2]

[1]National Research Institute of Fisheries Science, Fisheries Research Agency, 2-12-4 Fukuura, Kanazawa, Yokohama, Kanagawa 236-8648, Japan

[2]Atmosphere and Ocean Research Institution, The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8564, Japan

ABSTRACT: Given the significance of planktonic copepods to marine food webs and biogeochemical cycles, information about their broad-scale community structure can enhance our understanding of the global marine ecosystem. We performed a metagenetic analysis of nuclear large subunit ribosomal DNA to determine spatial patterns in the community structure of epipelagic copepods in the tropical and subtropical Pacific. The metagenetic method was taxonomically comprehensive, avoided time-consuming morphological classification, and successfully characterized copepod communities at 19 stations in Kuroshio, the North Pacific subtropical gyre (NPSG), the eastern tropical Pacific, and the South Pacific subtropical gyre (SPSG). A total of 404 molecular operational taxonomic units (MOTUs) were obtained from 434 304 quality-filtered sequence reads with a 97% similarity threshold. Using community analyses of sequence reads and MOTU compositions, we identified a distinct copepod community shoreward of the Kuroshio Current. The MOTU compositions were similar between the oceanic side of Kuroshio and the subtropical gyres; however, on the basis of sequence reads, the copepod communities were clustered in 2 groups, equator–Kuroshio and subtropical gyre. These groups were highly correlated with chl *a* concentrations, and primary productivity appeared to significantly determine the major distribution patterns of copepods in the tropical and subtropical Pacific. Multiple taxonomic groups co-existed in the oligotrophic subtropical gyres, where MOTU numbers and diversity indices were high, particularly in the NPSG. In addition to the large number of species with distribution peaks in the subtropical gyres, the relatively high productivity in the NPSG compared with the SPSG might favor some species with distribution peaks in the eastern tropical Pacific and Kuroshio regions, leading to the highest copepod diversity in the NPSG.

KEY WORDS: Copepoda · Metagenetics · Community structure · Diversity · Pacific Ocean

INTRODUCTION

Planktonic copepods are an ecologically important group within marine zooplankton, playing a significant role in marine food webs and in biogeochemical cycles (Mauchline 1998, Turner 2004). Copepods are ubiquitous in the oceans and are a highly abundant and diverse group. Copepods are sensitive to environmental changes and are thus considered indicators of natural and anthropogenic stressors (Roemmich & McGowan 1995, Hays et al. 2005). Changes in copepod community structure affect higher trophic levels and biogeographic shifts in copepods may correspond to reductions in the populations of commercially important fish (Beaugrand et al. 2003). Ocean environments and ecosystems are rapidly changing and experiencing increasing temperatures, acidification, overfishing, pollution, and introduction of non-indigenous species, among other stresses. As copepods can respond rapidly to changes in ocean

*Corresponding author: jhirai@affrc.go.jp

environments, an understanding of their community structure can help us detect and understand changes in marine environments and ecosystems. Therefore, it is important to reveal and characterize the current community structure of copepods globally.

The community structures of oceanic zooplankton taxa, including copepods, have been investigated in various regions. For example, large-scale surveys using continuous plankton recorders have been extensively conducted for more than 50 yr in the North Atlantic and have been developed more recently for the North Pacific and the Southern Ocean (Batten et al. 2003, Richardson et al. 2006). In the tropical and subtropical oceans, community analyses of zooplankton covering large areas have been performed in the Atlantic (Woodd-Walker et al. 2002, Schnack-Schiel et al. 2010). However, investigations in the Pacific have been restricted to local areas, such as the ocean areas near Japan, Hawaii, and the eastern tropical Pacific (Kidachi & Itoh 1979, McGowan & Walker 1979, Longhurst 1985, Landry et al. 2001). McGowan (1971, 1974) investigated the distribution of major planktonic species in the Pacific and provided significant insights into the biogeography of Pacific zooplankton taxa. However, little is known about the whole community structure of copepods and relationships among communities across a large area of the tropical and subtropical Pacific.

The inaccessibility of oceanic areas in the tropical and subtropical Pacific and the difficulty of morphological classification of these species across broad oceanic areas have constrained large-scale community studies of copepods. Morphological classification of copepod specimens is time-consuming and requires sophisticated expertise, particularly in tropical and subtropical areas, where species richness is high, specimens are small in size, and there is limited morphological information regarding copepods (Rombouts et al. 2010, Schnack-Schiel et al. 2010). Molecular techniques, however, can identify copepods on the basis of species-specific DNA sequences without relying on morphological characteristics, as well as identifying immature copepods and cryptic species (Bucklin et al. 2011). In addition, results obtained by molecular techniques may be compared easily, whereas the accuracy of data obtained by morphological classifications largely depends on the taxonomist.

The metagenetic approach, also called metabarcoding, reveals community structure and diversity from environmental DNA using a specific molecular marker. This technique has rapidly developed since the advent of high-throughput sequencing (Margulies et al. 2005). In the metagenetic approach, a high number of sequence reads are obtained from environmental samples, and community structure is investigated on the basis of molecular operational taxonomic units (MOTUs), which are genetic groups determined by sequence similarities in a specific molecular marker. The metagenetic approach has been developed primarily for microbial community analysis and has been used to evaluate the community structure of metazoans (Fonseca et al. 2010, Hajibabaei et al. 2011, Ji et al. 2013), including zooplankton (Lindeque et al. 2013, Pearman et al. 2014) and planktonic copepods (Hirai et al. 2015). Common molecular markers for metagenetic analysis of metazoans are nuclear small subunit ribosomal DNA (SSU rDNA) and mitochondrial cytochrome *c* oxidase subunit I (COI). However, nuclear large subunit ribosomal DNA (LSU rDNA) allows easier design of primer pairs than COI and analysis of larger genetic variations than SSU rDNA in copepods. Thus, LSU rDNA enables a rapid and taxonomically comprehensive metagenetic analysis of copepods without individual morphological identifications of specimens (Hirai et al. 2015). Although there is a smaller reference database available for the LSU rDNA region than for SSU rDNA or COI, the metagenetic method using the LSU rDNA region remains an effective tool for revealing the community structure of copepods, especially in the tropical and subtropical Pacific, where morphological classifications are difficult.

In this study, we revealed spatial patterns of copepod community structure in the epipelagic layer in the tropical and subtropical Pacific, using metagenetic analysis of the LSU rDNA region. Samples were collected in 4 ocean regions: Kuroshio, the North Pacific subtropical gyre (NPSG), the eastern tropical Pacific (ETP), and the South Pacific subtropical gyre (SPSG). First, we clustered copepod communities into major groups on the basis of sequence reads of MOTUs. Each clustered group was characterized by dominant taxa, and mechanisms covering community structure were investigated in relation to environmental variables. We also performed community analysis of MOTU compositions and investigated diversity patterns of copepods across the tropical and subtropical Pacific. To date, few community studies of epipelagic copepods have been performed in the tropical and subtropical Pacific, and little is known about the relationships of copepod communities across broad-scale oceanic regions. This study provides a basic understanding of the structures of copepod communities in the tropical and subtropical Pacific.

MATERIALS AND METHODS

Sampling

A total of 20 plankton samples were collected at 19 stations in 4 ocean regions in the tropical and subtropical Pacific: Kuroshio (K1–K3), the NPSG (N1–N8), the ETP (E1–E3), and the SPSG (S1–S5) (Fig. 1). Details regarding collection sites and dates are provided in Table S1 in Supplement 1 at www.int-res.com/articles/suppl/m534p065_supp.pdf. For this study, samples were obtained from 16 stations during the KH-11-10 and KH-12-01 cruises (Fig. 1) aboard the RV 'Hakuho-Maru' (Japan Agency for Marine–Earth Science and Technology). Samples were collected at 0–200 m at night using a vertical multiple plankton sampler (VMPS) (Terazaki & Tomatsu 1997) with a 0.5 m^2 mouth opening and 100 µm mesh. At Stn E1, samples were taken during both the day (E1D) and night (E1N). The data in our previously published study (Hirai et al. 2015) were also used for Stns K1–K3 in Kuroshio during the SY-11-05 cruise aboard the FRV 'Soyo-Maru' (National Research Institute of Fisheries Science). We collected these samples during the daytime using a VMPS with a 0.25 m^2 mouth opening. All samples were preserved in 99% ethanol and maintained at 4°C. Ethanol was replaced 24 h following the initial preservation.

Environmental variables were measured during the cruises. Temperature and salinity were obtained using a CTD system (SBE-911 plus, Sea-Bird Electronics), and dissolved oxygen was measured using an SBE-43 dissolved oxygen sensor (Sea-Bird Electronics). Water samples were collected with Niskin bottles attached to the CTD system and were filtered using Whatman GF/F filters for chl *a* analysis. Chl *a* was extracted with N,N-dimethylformamide, and the chl *a* concentration was analyzed using a Turner fluorometer (Welschmeyer 1994). Mixed layer depth (MLD) was calculated by the depth at a temperature of $\Delta T = 0.2$ °C from the temperature at 10 m depth (de Boyer Montégut et al. 2004). Vertical profiles of temperature, salinity, and chl *a* are provided in Supplement 1 (Figs. S1–S3).

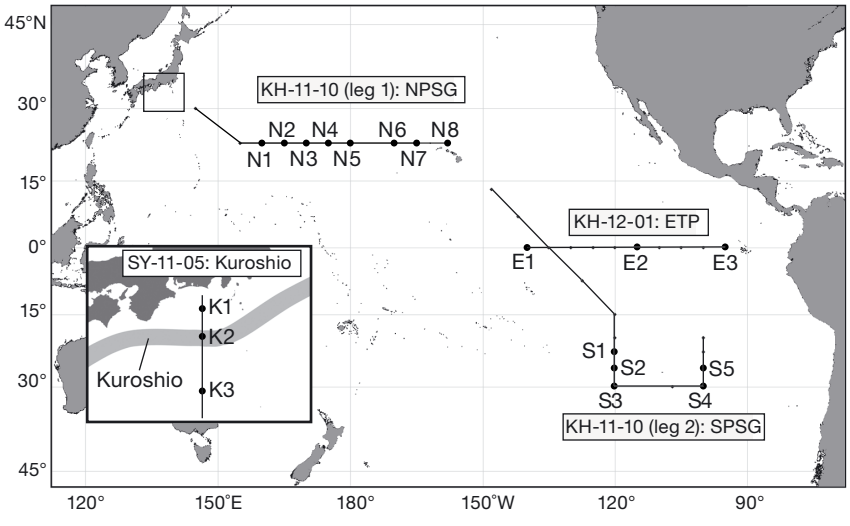

Fig. 1. Sampling locations and cruise tracks. Zooplankton samples were collected at 19 stations during 3 cruises (SY-11-05, KH-11-10, and KH-12-01) in 4 ocean regions: Kuroshio (Stns K1–K3), the North Pacific subtropical gyre (NPSG; Stns N1–N8), the eastern tropical Pacific (ETP; Stns E1–E3), and the South Pacific subtropical gyre (SPSG; Stns S1–S5). Samples were collected both during the day and at night at Stn E1

DNA extraction, PCR amplification, and 454 pyrosequencing

DNA extraction, PCR amplification, and 454 pyrosequencing were mostly performed according to Hirai et al. (2015). Total genomic DNA was extracted from newly collected samples using the Gentra Puregene Cell and Tissue Kit (QIAGEN); in the first step, 12 ml of Cell Lysis Solution was added to a 50 ml tube and samples were incubated for 3 h at 65°C. Three ethanol-preserved samples from Stns E1–E3 were split, and due to the large amount of biomass in these samples, half aliquots were used for DNA extraction. DNA extraction was performed for all samples within a year after sample collection. The LSU rDNA D2 region was amplified using primers LSU Cop-D2F (5′-AGA CCG ATA GCA AAC AAG TAC-3′) and LSU Cop-D2R (5′-GTC CGT GTT TCA AGA CGG-3′). PCR reactions were performed in a 25 µl volume containing 8 µl of distilled water, 12.5 µl of 2× Gflex PCR buffer (Takara), 1 µl of each primer (5 µM), 0.5 µl of Tks Gflex DNA Polymerase (Takara), and 2 µl of template DNA (1 ng μl^{-1}). PCR cycling included initial denaturation at 94°C for 1 min, followed by 22 cycles of 10 s denaturation at 98°C, 15 s annealing at 58°C, and 1 min extension at 68°C. A final extension step was performed at 68°C for 7 min. A 1 µg volume of each product was purified with the MinElute PCR purification kit (QIAGEN), and 3 pyrosequencing runs (quarter-plate runs) were carried out on a 454 GS FLX system (Roche). Each sample had a unique

multiplex identifier (MID) adaptor from the GS Titanium Rapid Library MID Adaptors Kit (Roche). The MID tags used for the samples are listed in Table S1. All procedures for 454 pyrosequencing were performed according to the manufacturers protocol. Raw 454 data files are available in the NCBI/EBI/DDBJ Sequence Read Archive: DRA002161 and DRA002227.

MOTU clustering

Data analysis of raw reads was performed using mothur (Schloss et al. 2009). The newly obtained 454 data were combined with previously obtained data from K1–K3, all sequence reads were quality-filtered, and copepod sequences were selected, mainly according to the method in Hirai et al. (2015). In order to align sequence reads, we used the add-fragments option in Multiple Alignment using Fast Fourier Transform (MAFFT) with the default setting (Katoh & Standley 2013). The reference sequences for alignment for MAFFT are available as a supplementary file (Supplement 2 at www.int-res.com/articles/suppl/m534p065_supp.xls), which contains LSU rDNA D2 sequences of copepods obtained from GenBank (excluding environmental sequences). The aligned sequences were filtered using single-linkage preclustering (Huse et al. 2010). Possible chimeras were removed by UCHIME without a reference dataset (Edgar et al. 2011).

MOTU clustering was performed for all quality-filtered sequences at a 97% similarity threshold, which was a suitable level for avoiding artificial inflation of MOTUs using 454-pyrosequencing and for characterizing the community structure of pelagic copepods using LSU rDNA regions (Kunin et al. 2010, Hirai et al. 2015). We used only MOTUs containing ≥3 sequence reads to avoid overestimating MOTUs (Hirai et al. 2015). The number of MOTUs was used as a proxy for species richness. The use of sequence reads as quantitative data is sometimes avoided due to methodological biases in metagenetic analysis (Amend et al. 2010). However, sequence reads tended to increase with increasing biomass in zooplankton (Lindeque et al. 2013). Although metagenetic analysis of planktonic copepods is not free from potential bias (e.g. amplification efficiency of universal primers), this method effectively detects dominant taxa in community analysis (Hirai et al. 2015). Therefore, we used sequence reads as an informative proxy for biomass in this study.

MOTUs were classified into the Orders Calanoida, Cyclopoida, Harpacticoida, and Poecilostomatoida using the naive Bayesian classifier (Wang et al. 2007) in mothur, with a 70% threshold; taxonomic information was derived from Boxshall and Halsey (2004). MOTUs below the 70% threshold were treated as 'unclassified' MOTUs. The order Calanoida is a well-studied group with relatively large numbers of registered LSU rDNA sequences, and calanoid MOTUs were classified as a family. The genus *Calocalanus* was further distinguished from the family Paracalanidae due to its importance in community analysis. Calanoid MOTUs may be classified successfully into taxonomic families using the classification method in this study (see Hirai et al. 2015), and we confirmed that the genus *Calocalanus* is phylogenetically distinct in LSU rDNA sequences from other genera in the Family Paracalanidae. To standardize the sequence data, the minimum number of sequence reads among all samples was randomly selected from each sample in mothur, and these standardized sequence reads were used for the following community and diversity analyses.

Community analysis based on sequence reads

Community analysis of copepods was conducted using sequence reads. Similarities within the copepod community were investigated using Bray-Curtis similarity by cluster analysis using the group average and by multidimensional scaling (MDS). Cluster and MDS analysis were performed in PRIMER v.6 (Clarke & Warwick 2001). Previous studies of copepods that used morphological classifications were conducted on the basis of the abundance (number of individuals); however, metagenetic community analysis is based on proportions of sequence reads. Proportions of sequence reads in each taxonomic group were compared between stations.

The relationship between copepod community (sequence reads) and environmental variables (water temperature, salinity, dissolved oxygen, chl *a*, and MLD) was assessed using redundancy analysis (RDA) in CANOCO v.5, to detect environmental factors related to dominant MOTUs in sequence reads. RDA was selected due to the result of detrended canonical correspondence analysis. We used both sea surface temperature (SST) and depth-averaged (0–200 m) temperature for water temperature parameters. Salinity and oxygen measures were also averaged across 0–200 m, and we used integrated chl *a* concentration (mg m^{-2}) at 0–200 m. The forward selection method (ter Braak & Verdonschot 1995) was used to determine and order statistically significant

environmental variables for RDA ($p < 0.05$). A Monte Carlo permutation test was performed using 499 unrestricted permutations to test whether environmental variables were significantly correlated with the copepod community composition. The significance of the first axis and all combined axes in RDA ordination were also tested by 499 permutations.

We also performed cluster analysis for major MOTUs, which were defined as those with at least 30 sequence reads in a single sample. This criterion retained over 20% of total MOTUs, which was adequate for detecting major distribution patterns. To standardize the sequence data of each MOTU, sequence reads were divided by total sequence reads. Therefore, this cluster analysis was based on the relative distribution peaks of sequence reads in each MOTU.

Spatial patterns of MOTU compositions and diversity

Community analysis of copepods was also conducted using MOTU compositions (presence/absence of MOTU). Both cluster and MDS analyses were performed using MOTU compositions, as described in the previous section. The numbers of MOTUs in each taxonomic group were also compared between stations.

Copepod diversity at each station was evaluated using the total number of MOTUs, the Shannon-Wiener diversity index (H'), and the Simpson diversity index (D). Non-parametric Kruskal-Wallis tests were used to identify significant differences in observed MOTU numbers among ocean regions. A significant difference between 2 ocean regions was tested using a non-parametric Mann-Whitney U test. Correlations between environmental variables and MOTU richness and diversity indexes were determined using Pearson's product-moment correlation coefficients (r). Statistical analyses were performed using SPSS 21.0. Numbers of shared MOTUs among ocean regions were also estimated using equal numbers of sequence reads. Only night samples were used from Stn E1 for statistical analyses and estimated numbers of shared MOTUs.

RESULTS

Community analysis based on sequence reads

A total of 434 304 quality-filtered sequence reads were obtained, and clustered into 404 MOTUs at a 97% similarity threshold. A total of 230 MOTUs, including 366 266 sequence reads, were classified into calanoid copepods. Randomly selected sequence reads (n = 6229) from each sample were standardized and used for the analyses. This standardized dataset included 383 of 404 MOTUs.

Copepod community compositions determined by sequence reads were strongly affected by ocean areas and clustered into 3 major groups: (1) Stn K1, (2) the equator–Kuroshio group containing Stns K2–K3 and the ETP stations (E1–E3), and (3) the subtropical gyre group, containing Stns N1–N8 and S1–S5 (Fig. 2a). Samples collected during the day and night at Stn E1 showed the highest similarity of all samples. The community structure at Stn S4 was distinct from other samples in subtropical gyres. The NPSG was more similar in community structure than the SPSG was to the equator–Kuroshio group (Fig. 2b).

High proportions of Calanidae, Clausocalanidae, Paracalanidae, and Cyclopoida sequence reads were observed across stations (Fig. 2c). In addition to these dominant taxa, Eucalanidae was also dominant in Group 1 of the cluster analysis (Stn K1). Other taxa only accounted for small proportions of sequence reads in Group 1. Dominant taxa were similar between Group 1 and Group 2 (the equator–Kuroshio group), however, proportions of sequence reads increased in various taxa (e.g. Clausocalanidae and Euchaetidae) in Group 2. In Group 2, we also observed taxa that were absent in Group 1 (e.g. Augaptilidae, *Calocalanus*, Centropagidae). The co-existence of various taxa was especially evident in Group 3 (the subtropical gyre group). Compared with Groups 1 and 2, higher proportions of sequence reads were observed in Group 3 in various taxa, such as Aetideidae, Augaptilidae, *Calocalanus*, Candaciidae, Centropagidae, Lucicutiidae, Metridinidae, and Poecilostomatoida. However, no Rhincalanidae and only a small proportion of Eucalanidae were observed in Group 3.

RDA ordination

RDA ordination was applied to investigate environmental factors affecting community structure based on sequence reads. Forward selection methods revealed that chl *a* was the most significant environmental variable explaining copepod community structure in the first axis and all axes combined (F = 6.0, p = 0.002), followed by SST (F = 4.0, p = 0.002), dissolved oxygen (F = 2.9, p = 0.024), and MLD (F = 2.3, p = 0.036). These 4 environmental variables explained 30.9% of the variance of MOTU data by the

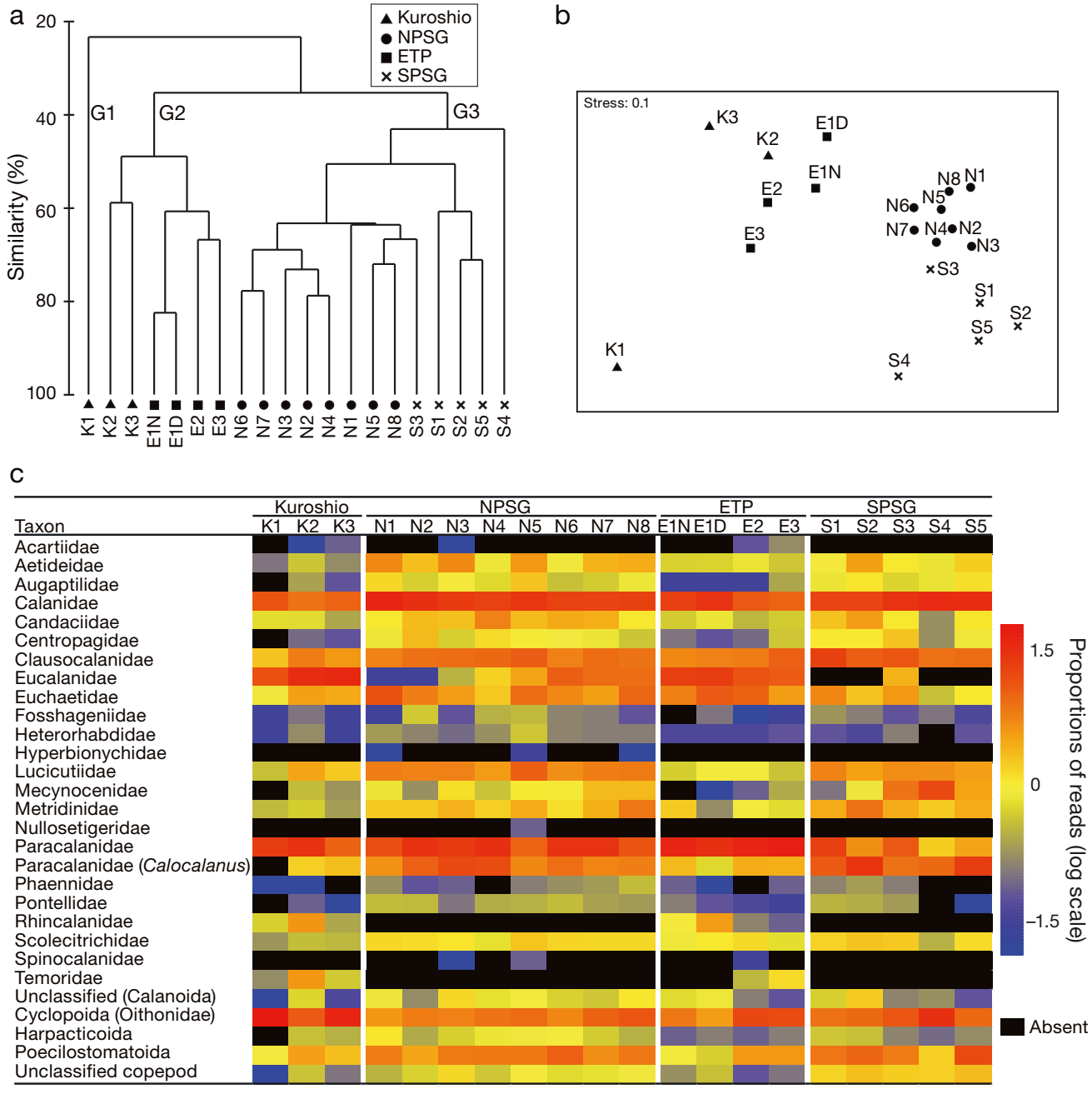

Fig. 2. Community analyses of sequence read data showing similarities within copepod communities. (a) Cluster analysis dendrogram showing 3 copepod community groupings G1–G3, (b) multidimensional scaling ordination, and (c) taxonomic compositions of copepod sequence reads for samples collected from ocean regions of Kuroshio (Stns K1–K3), the North Pacific subtropical gyre (NPSG; Stns N1–N8), the eastern tropical Pacific (ETP; Stns E1–E3), and the South Pacific subtropical gyre (SPSG; Stns S1–S5). Proportions of sequence reads in each taxon were log transformed at each station to enable comparisons of dominant taxa

first axis and 13.6% by the second axis in RDA ordination (Fig. 3). The variability of the copepod community structure was explained significantly by both the first axis and all the axes combined ($p < 0.01$). The 3 clustered groups (Groups 1–3) obtained in the cluster analysis (Fig. 2a) were clearly separated by the environmental variable chl *a*. Group 1 was also distinguished from other samples mostly by its low SST value. Additionally, the factors SST and dissolved oxygen differed between Kuroshio and ETP.

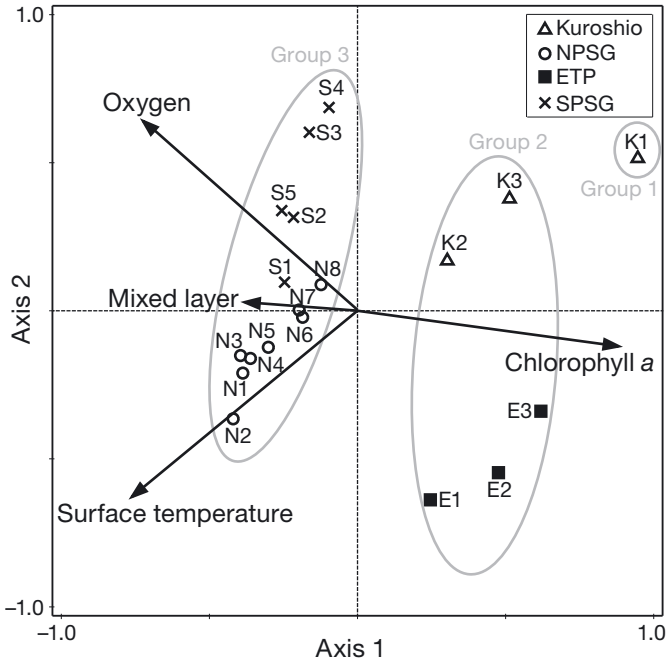

Fig. 3. Redundancy analysis ordination for copepod communities and environmental variables sampled across ocean regions of Kuroshio (Stns K1–K3), the North Pacific subtropical gyre (NPSG; Stns N1–N8), the eastern tropical Pacific (ETP; Stns E1–E3), and the South Pacific subtropical gyre (SPSG; Stns S1–S5). Black arrows indicate significant environmental variables for molecular operational taxonomic units (MOTUs) data using the forward selection method. A total of 30.9 and 13.6% of the variance of MOTU data were explained by the first and second axes, respectively. The arrow length of environmental variables indicates the strength of its relationship with community structures. Gray circles represent the 3 major groupings determined from cluster analysis (see Fig. 2a)

Distribution patterns of MOTUs

Cluster analysis of the 83 major MOTUs revealed 7 groups (A–G) at the 40% similarity level based on the relative distribution peaks of sequence reads (Fig. 4a,b). The majority of MOTUs were clustered into Groups C (47/83 MOTUs) and D (17/83 MOTUs). Both groups were widely distributed, however, the distribution peaks of Group C occurred in subtropical gyres, whereas those of Group D in the ETP and Kuroshio regions (Fig. 4c). Other groups showed distinct distribution peaks in specific stations: Stns N6–N8 in Group A, Stn S4 in Group B, Stn K1 in Group E, Stn E1 in Group F, and Stn K2 in Group G. The clusters were related to the taxonomy of MOTUs (Fig. 4a). For example, MOTUs in genus *Calocalanus* (8/83 MOTUs) were only observed in subtropical gyre Groups B and C, whereas no MOTUs in Eucalanidae (7/83 MOTUs) were present in these groups. Temoriidae MOTUs were restricted to Group G and Rhincalanidae MOTUs to Groups F and G. Meanwhile, various distribution patterns of MOTUs were observed in taxonomic groups of Paracalanidae, except for *Calocalanus* (7 MOTUs in Groups A, C, and D), Cyclopoida (9 MOTUs in Groups C–E), and Poecilostomatoida (13 MOTUs in Groups A, C, and D). Various taxa of 4 orders and 11 families in calanoid copepods were present in Group C with distribution peaks in subtropical gyres.

Community analysis based on MOTU compositions

Community structure differed strongly between ocean areas according to both MOTU composition, as well as of community analysis using sequence reads (Fig. 5a,b). However, copepod communities identified from MOTU compositions showed that communities in the oceanic side of Kuroshio (i.e. Stns K2 and K3) were more similar to those in subtropical gyres than to those in ETP (Fig. 5a). MOTU compositions in the oceanic side of Kuroshio or the ETP were more similar to those in the NPSG than to those in the SPSG (Fig. 5b). The differences in the MOTU number of each taxon showed that most of the families had relatively large MOTU numbers in the subtropical gyre, particularly in the NPSG (Fig. 5c), except for a few specific taxa, such as Eucalanidae, Rhincalanidae, and Temoriidae, that possessed a larger number of MOTUs in the ETP and Kuroshio regions than in subtropical gyres. At Stn K1, some taxa were not present (or fell below the threshold value), including the families Augaptilidae, Centropagidae, Mecynoceridae, and Pontellidae and the genus *Calocalanus*.

Spatial patterns of copepod diversity

The observed MOTU number and diversity index (H' and $1–D$) were higher in NPSG than other ocean regions, with the highest number of observed MOTUs (166) sampled at Stn N3 (Table 1). MOTU number and diversity index were also high in the SPSG, although those values decreased at Stn S4. The diversity index in the ETP was low compared with those in NPSG and SPSG (except for Stn S4). MOTU number was also low at Stns E1 and E3; however, Stn E2 showed high MOTU richness that was almost equivalent or superior to that observed in the SPSG. At Stn E1, the MOTU number increased 11%

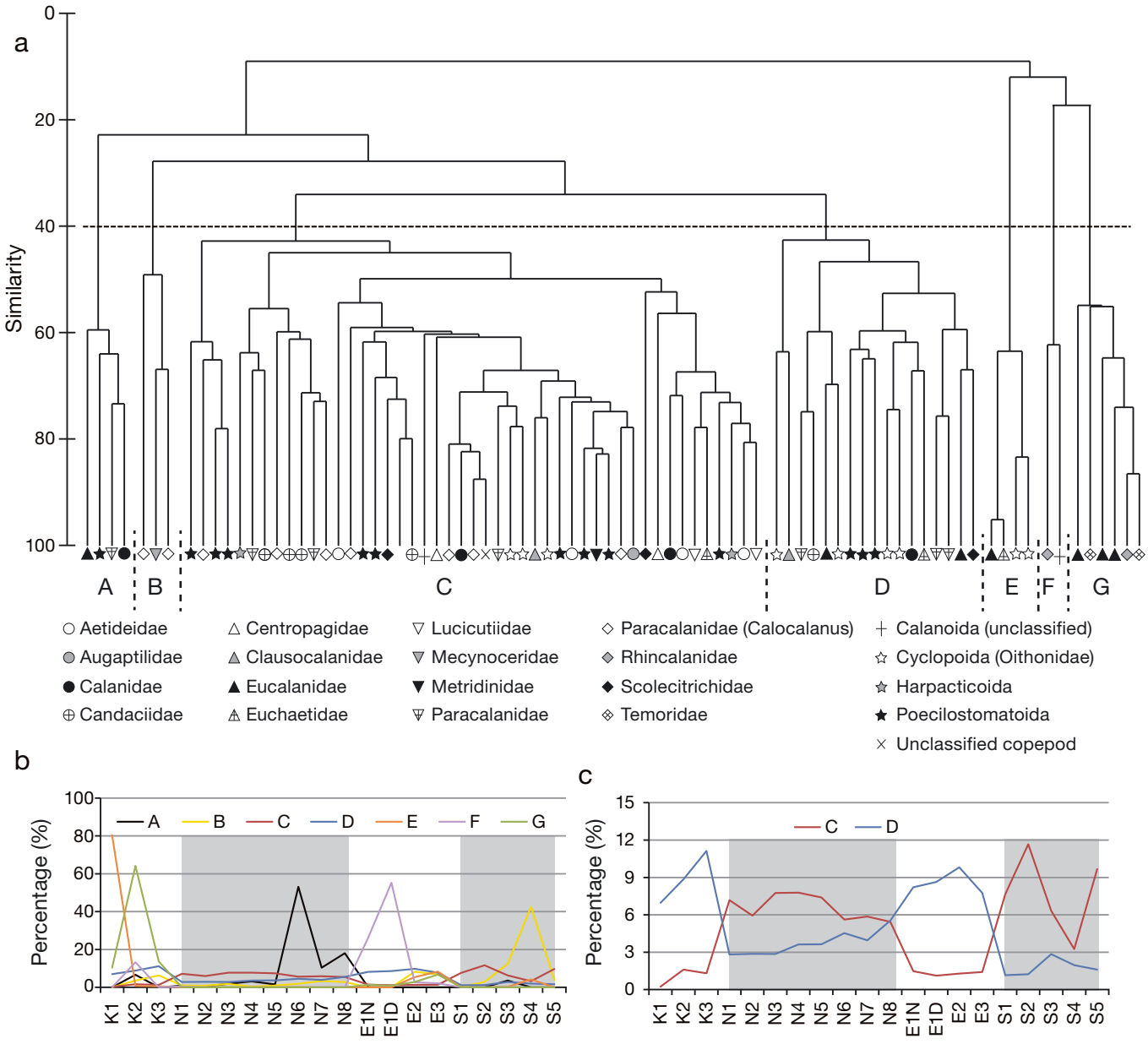

Fig. 4. (a) Cluster analysis of major molecular operational taxonomic units (MOTUs), (b) average relative distribution patterns showing peaks of 7 clustered groups (A–G) at the 40% similarity level, and (c) distribution patterns of the major Groups C and D. A total of 83 MOTUs with at least 30 sequence reads in a single site were clustered by their relative peaks of distribution. Gray areas indicate Stns N1–N8 in the North Pacific subtropical gyre (NPSG) and S1–S5 in the South Pacific subtropical gyre (SPSG), white areas indicate Stns K1–K3 in Kuroshio and Stns E1–E3 in the eastern tropical Pacific (ETP)

among samples collected at night, and diversity indices were higher for night-collected samples than for those collected during the day. The lowest MOTU number and diversity indexes were observed at Stn K1 in Kuroshio. Within Kuroshio, MOTU number and diversity index peaked at Stn K2, with a higher diversity index than in the ETP.

Observed numbers of MOTUs were significantly different among ocean regions (Kruskal-Wallis; $p < 0.005$), and the number of observed MOTUs was significantly higher in the NPSG than in other areas (Mann-Whitney U; $p < 0.05$). All MOTU numbers and diversity indices were correlated positively with SST ($p < 0.01$) and negatively with chl *a* ($p < 0.01$) (Table 2). MOTU number was most correlated with SST ($r = 0.77$). In contrast, average temperature was not correlated with MOTU number. Diversity indices were most strongly (negative) correlated with total chl *a* ($r = -0.762$ and -0.792 for H' and D, respectively).

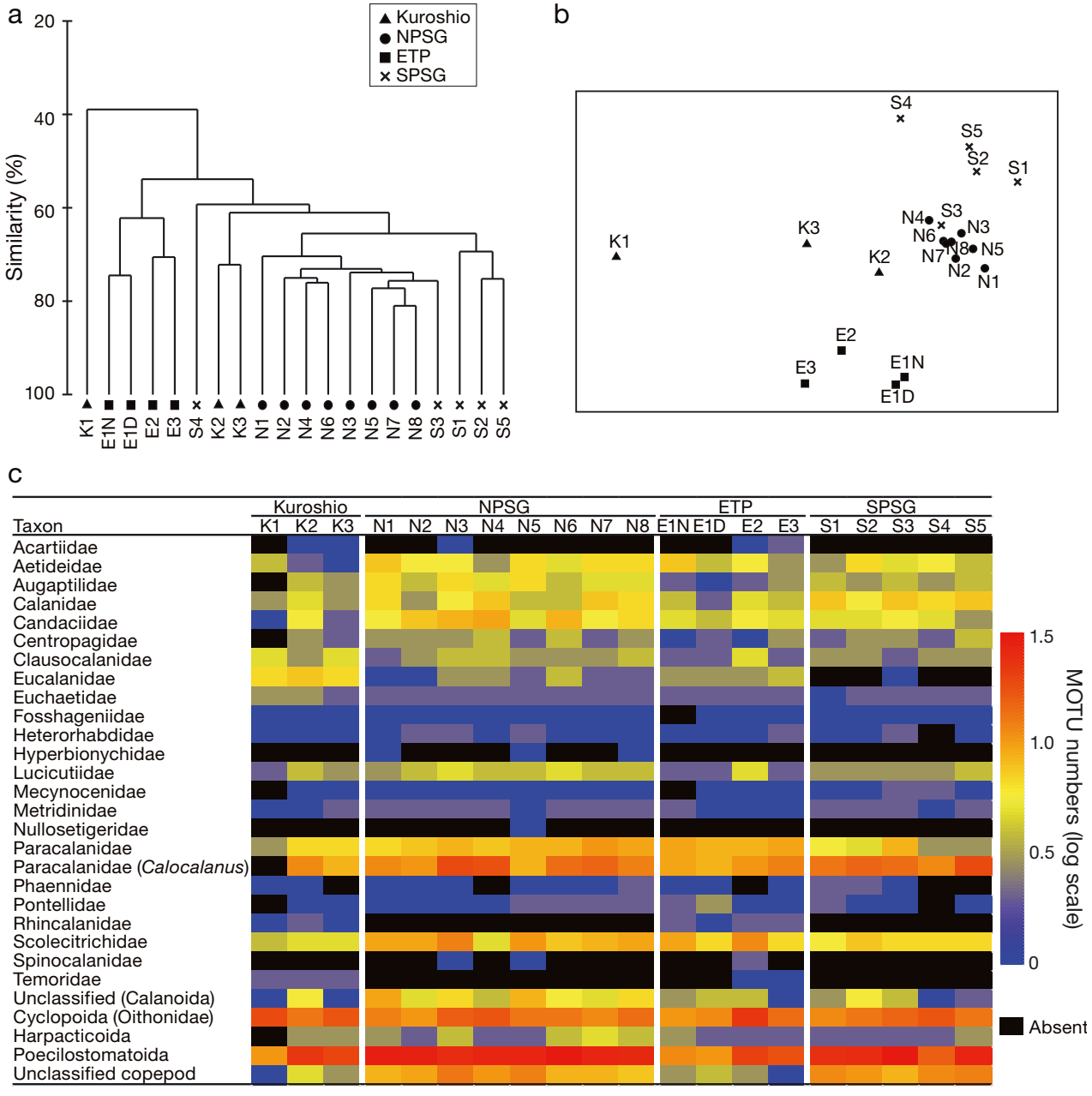

Fig. 5. Community analysis based on molecular operational taxonomic unit (MOTU) compositions. (a) Cluster analysis dendrogram, (b) multidimensional scaling ordination, and (c) taxonomic compositions of copepod MOTUs for samples collected from ocean regions of Kuroshio (Stns K1–K3), the North Pacific subtropical gyre (NPSG; Stns N1–N8), the eastern tropical Pacific (ETP; Stns E1–E3), and the South Pacific subtropical gyre (SPSG; Stns S1–S5). MOTU numbers in each taxon were log transformed at each station

From 18 687 randomly selected reads in each ocean region, the number of observed MOTUs was highest in the NPSG (247 MOTUs), followed by the SPSG (204 MOTUs), Kuroshio (163 MOTUs), and the ETP (176 MOTUs) (Table 3). A total of 84 MOTUs were shared between all areas. The MOTUs in the NPSG were also frequently present in other ocean regions, with 128 of 163 (78.5%) MOTUs in Kuroshio and 121 of 176 (68.8%) MOTUs in the ETP also observed in the NPSG. The highest number of shared MOTUs (159) was shared between the NPSG and the SPSG, accounting for 64.3% of MOTUs in the NPSG and 77.9% of MOTUs in the SPSG.

Table 1. Molecular operational taxonomic unit (MOTU) number and diversity index, Shannon-Wiener (H') and Simpson (1–D), for each sampling station. Standardized sequence data (6229 sequence reads) were used for each sample. NPSG, ETP and SPSG denote North Pacific subtropical gyre, eastern tropical Pacific, and South Pacific subtropical gyre, respectively

Area	Station	MOTUs	H'	1–D
Kuroshio	K1	72	2.11	0.78
	K2	125	3.25	0.92
	K3	105	2.62	0.84
NPSG	N1	143	3.18	0.91
	N2	130	3.19	0.92
	N3	166	3.60	0.95
	N4	145	3.44	0.93
	N5	143	3.60	0.95
	N6	147	3.56	0.95
	N7	143	3.48	0.94
	N8	146	3.62	0.96
ETP	E1 (night)	99	2.80	0.90
	E1 (day)	89	2.65	0.89
	E2	133	2.82	0.88
	E3	108	2.84	0.90
SPSG	S1	119	3.19	0.92
	S2	129	3.58	0.95
	S3	139	3.52	0.95
	S4	107	2.73	0.89
	S5	129	3.36	0.93

Table 2. Pearson's correlations between environmental variables and the copepod community molecular operational taxonomic unit (MOTU) number and diversity indexes (Shannon-Wiener [H'] and Simpson [1–D]). $^{*}p < 0.05$, $^{**}p < 0.01$

	MOTUs	H'	1–D
Surface temperature	0.77**	0.676**	0.746**
Average temperature	0.44	0.699**	0.641**
Salinity	0.308	0.465*	0.524*
Chl *a*	–0.684**	–0.762**	–0.792**
Dissolved oxygen	0.171	0.519*	0.412
Mixed layer depth	0.77	0.325	0.169

Table 3. Number of shared molecular operational taxonomic units (MOTUs) according to ocean regions Kuroshiro, North Pacific subtropical gyre (NPSG), the eastern tropical Pacific (ETP), and South Pacific subtropical gyre (SPSG). The standardized sequence data (18 687 reads) in each ocean region were used to calculate observed and shared numbers of MOTUs

	MOTUs	Shared MOTUs			
		Kuroshio	NPSG	ETP	SPSG
Kuroshio	163		128 (78.5%)	108 (66.3%)	113 (69.3%)
NPSG	247	128 (51.8%)		121 (49.0%)	159 (64.3%)
ETP	176	108 (61.3%)	121 (68.8%)		102 (58.0%)
SPSG	204	113 (55.4%)	159 (77.9%)	102 (50.0%)	

DISCUSSION

Community structure of copepods

As global ocean environments are rapidly changing, it is important to document the current community structure of copepods and reveal the mechanisms that govern spatial patterns in the copepod community. In this study, a metagenetic analysis using the LSU rDNA region successfully revealed spatial patterns of community structure of epipelagic copepods in the tropical and subtropical Pacific, using both sequence reads and MOTU compositions. Hydrographic conditions were significantly different at the boundary of the Kuroshio Current (Figs. S1–S3 in Supplement 1), and a distinct copepod community was identified on the shoreward side of the current (Figs. 2 & 5). Other copepod communities were clearly separated by ocean areas, almost following the boundaries of Longhurst biogeographic provinces (Longhurst 2007), which suggests that copepod communities are strongly affected by different environmental conditions in water masses.

MOTU compositions of copepods in the oceanic side of the Kuroshio region (Stns K2 and K3) were clustered into the subtropical gyre group, comprising the NPSG and the SPSG (Fig. 5a). However, community analysis based on sequence reads showed high similarity between the oceanic side of the Kuroshio region and the ETP (Fig. 2a). We found strong relationships between chl *a* and the copepod community based on sequence reads, indicating that food conditions significantly influence the dominance of copepods in the tropical and subtropical Pacific. Strong relationships between primary production and dominant copepod taxa were reported from a large-scale community analysis in the Atlantic (Woodd-Walker et al. 2002). Therefore, species distributed in the Kuroshio region also appear in subtropical gyres; however, higher productivity in Kuroshio than in the subtropical gyres has led to the common appearance of certain dominant taxa between Kuroshio and the ETP. According to both analyses of sequence reads and MOTU compositions, copepod communities in the oligotrophic NPSG and SPSG were similar, indicating comparable species compositions and dominant taxa in these regions. This finding agrees with a previous study that reported highly similar community structure between the NPSG and

SPSG according to morphological analysis (Williamson & McGowan 2010).

While common MOTUs were found across many different stations and ocean regions, the relative number of reads indicated several distribution patterns of MOTUs. Cluster analysis of MOTU distributions revealed major distribution patterns in the ETP and Kuroshio regions and subtropical gyres groups (Fig. 4). These distribution patterns of zooplankton are common in the Pacific (McGowan 1971); however, we also detected MOTUs with distribution patterns in specific areas, including Stns N6–N8 in the eastern side of the NPSG (Group A), S4 in the SPSG (Group B), K1 in the shoreward region of the Kuroshio Current (Group E), E1 in the western side of the ETP (Group F), and K2 on the Kuroshio current (Group G) (Fig. 4). Although changes in environmental conditions and community structure of copepods are large between ocean areas, there are also environmental differences within ocean areas (Figs. S1–S3 in Supplement 1). Some of the observed oceanographic features, such as intrusion of cold water at Stn S4 in the SPSG and relatively high chl *a* concentrations at Stns N6–N8 in the NPSG, were associated with different metagenetic sample profiles. Therefore, some MOTUs are adapted to specific local environments within the complex species-specific biogeography of copepods in the tropical and subtropical Pacific.

In this study, some methodological variations occurred during sampling. VMPS nets with half-mouth opening areas were used during the daytime in Kuroshio, and divided samples were used for DNA extraction in Kuroshio and the ETP. Although methodological differences may cause a bias in metagenetic data, this bias may be alleviated by the relatively large biomass in Kuroshio and the ETP. The cluster analysis showed the smallest difference between samples collected by day and by night at Stn E1, additionally, Woodd-Walker et al. (2002) mentioned that sampling time was not a major source of difference in their large-scale community analysis of copepods. Therefore, methodological biases introduced by different sampling times are considered to be small, compared with community differences among sampling sites. Seasonal changes were not more evident at lower latitudes than at higher latitudes (Longhurst 2007), and seasonality was not considered in this study.

Characteristics of copepod communities

Spatial patterns of copepod communities were revealed and each community was also characterized by taxonomic information obtained by metagenetic analysis. Some taxa, such as Calanidae, Clausocalanidae, Paracalanidae, and Cyclopoida, were dominant across ocean regions (Fig. 2c), however, distribution peaks were MOTU-specific within these taxa (Fig. 4). Therefore, these taxa contained various species with specific environmental preferences (e.g. oligotrophic or eutrophic) associated with different distribution patterns in the tropical and subtropical Pacific. We also detected taxonomic groups that clearly characterized copepod communities. For example, the ETP and Kuroshio regions, including Stn K1, was characterized by dominant sequence reads of the family Eucalanidae. This family was dominant under high chl *a* conditions with MOTU-specific distribution patterns, suggesting that it was adapted to high food availability. In oligotrophic subtropical gyres, Eucalanidae were present as a small proportion of total species or restricted to relatively high chl *a* areas at Stns N6–N8. One of the dominant MOTUs at Stn K1 was identified by BLAST search as *Eucalanus californicus* (100% similarity) in the Family Eucalanidae (data not shown). This species maintains its population with ontogenetic vertical migration in areas of high primary production in the western North Pacific (Shimode et al. 2012). In the ETP, other species within Eucalanidae, including *Subeucalanus subtenuis* and *S. subcrassus*, have been reported as dominant species (Chen 1986). The distribution patterns of some taxa (e.g. Temoriidae, Rhincalanidae) were restricted to specific locations in Kuroshio or the ETP, and these MOTUs might be adapted to specific areas such as the core of the Kuroshio Current.

Compared with the ETP and Kuroshio regions, the high proportions of sequence reads and large numbers of MOTUs belonging to many different taxa were characteristics of communities in the subtropical gyres (Figs. 2c & 5c). The co-existence of multiple taxa may be an adaptation to oligotrophic areas where food sources are poor (McGowan & Walker 1979). A community structure with few dominant taxa in oligotrophic areas is also indicated by the negative correlation between the diversity index and levels of chl *a* (Table 2), which differentiates communities in the subtropical gyres from those in the ETP and Kuroshio regions. Some taxa are clearly adapted to the oligotrophic oceans; for example, the genus *Calocalanus* was particularly diverse and its proportion of sequence reads was higher in low chl *a* areas in this study. In previous work in the Atlantic, the genus *Calocalanus* was abundant in subtropical areas where low chl *a* was observed (Longhurst 2007, Schnack-Schiel et al. 2010).

Diversity patterns of copepods

Metagenetic analysis revealed the diversity pattern of copepods in the tropical and subtropical Pacific, and MOTU compositions were compared between stations. MOTU number was most correlated with SST, and high species richness was observed in subtropical gyres, particularly in the NPSG, where the highest water temperature was observed. Rombouts et al. (2010) also reported that SST was the environmental factor that best explained species richness of copepods; high species richness is predicted for conditions of high temperature and salinity, and low chl *a*, all of which are characteristics of subtropical gyres in both the Pacific and Atlantic Oceans. The finding of highest species richness in subtropical gyres is consistent with findings from previous studies of oceanic taxa, including those on copepods in the Atlantic (Angel 1993, Rombouts et al. 2009). Our results suggested that more copepod species are present in the NPSG than in the SPSG, which agrees with a report by Williamson and McGowan (2010). In the present study, the high species richness in the NPSG was explained by large numbers of MOTUs shared with other ocean areas. In addition to the large number of subtropical gyre species distributed in the NPSG and the SPSG, the NPSG contained more MOTUs that were also distributed in the ETP and Kuroshio regions than did the SPSG (Table 3). Community analyses based on MOTU compositions also showed that the species compositions on the oceanic side of the Kuroshio and the ETP regions were more similar to compositions in the NPSG than to those in the SPSG (Fig. 5b).

One possible explanation for high diversity in the NPSG is the strong Kuroshio Current and its recirculation, which transports species with distribution peaks in the ETP and Kuroshio regions to the NPSG. Population genetics of copepods show that haplotypes are highly shared between Kuroshio and the NPSG (Norton & Goetze 2013), suggesting that the NPSG and Kuroshio Current regions are highly connected. In addition, higher chl *a* concentration in the NPSG compared with the SPSG might favor species with distribution peaks in the ETP and Kuroshio regions that are adapted to take advantage of high food availability. In a comparison between the NPSG and the SPSG, the proportion of MOTUs with distribution peaks in the ETP and Kuroshio regions (see Group D in Fig. 5c) was higher in the NPSG (mean = 3.73%) than in the SPSG (mean = 1.76%), suggesting that species with distribution peaks in the ETP and Kuroshio regions are more adapted to the NPSG than to the SPSG. Zooplankton diversity is known to peak at the intermediate level of zooplankton biomass (Irigoien et al. 2004). The SPSG in this study area is the most oligotrophic area in the world, which might explain its smaller zooplankton biomass and lower diversity of copepods compared with the NPSG.

CONCLUSIONS

This is the first study to investigate the large-scale community structure of copepods in the tropical and subtropical Pacific using a metagenetic analysis. The analysis successfully revealed patterns of community structure and diversity of copepods in each ocean area, and the method was shown to be effective for a large-scale copepod community analysis. This study successfully characterized copepod communities and revealed the relationships among them. The community structure of copepods was unique in the Kuroshio region. Although MOTU composition in the oceanic side of Kuroshio was more similar to that of subtropical gyres, high productivity led to the dominance of MOTUs with a distribution peak in the ETP and Kuroshio regions. In addition, species with distribution peaks in the ETP and Kuroshio regions potentially extend into the NPSG, where productivity was higher than in the SPSG, and where the highest copepod diversity was found.

The method used in this study may easily be adapted for other field-collected zooplankton samples, and the data obtained in this study may be compared with those obtained in future analogous studies. The distribution patterns of MOTUs suggest that copepod biogeography is complex; additional sample collections across different temporal and spatial scales will likely reveal finer-scale community structure, diversity, and distribution patterns of copepods, thus improving our understanding of marine ecosystems and enabling us to better monitor environmental changes.

Acknowledgements. We thank the captains, crews, and researchers aboard the FRV Soyo-Maru and RV Hakuho-Maru for their assistance with field collections. We also thank the Center for Omics and Bioinformatics, Graduate School of Frontier Sciences, at The University of Tokyo, for the sequencing work and analysis. We thank S. Nishida, K. Hamasaki, K. Takahashi, S. Itoh (The University of Tokyo), and E. Goetze (University of Hawaii) for their constructive comments on an earlier draft of this paper. This work was supported by grants from the Japan Society for the Promotion of Science (grant nos. 247024 and 24121004).

LITERATURE CITED

Amend AS, Seifert KA, Bruns TD (2010) Quantifying microbial communities with 454 pyrosequencing: Does read abundance count? Mol Ecol 19:5555–5565

Angel MV (1993) Biodiversity of the pelagic ocean. Conserv Biol 7:760–772

Batten SD, Clark R, Flinkman J, Hays G, and others (2003) CPR sampling: the technical background, materials and methods, consistency and comparability. Prog Oceanogr 58:193–215

Beaugrand G, Brander KM, Lindley JA, Souissi S, Reid PC (2003) Plankton effect on cod recruitment in the North Sea. Nature 426:661–664

Boxshall GA, Halsey SH (2004) An introduction to copepod diversity. The Ray Society, London

Bucklin A, Steinke D, Blanco-Bercial L (2011) DNA barcoding of marine metazoa. Ann Rev Mar Sci 3:471–508

Chen YQ (1986) The vertical distribution of some pelagic copepods in the eastern tropical Pacific. CalCOFI Technical Report 27, California Cooperative Oceanic Fisheries Investigations, CA 27:205–227

Clarke KR, Warwick RM (2001) Change in marine communities: an approach to statistical analysis and interpretation, 2nd edn. PRIMER-E, Plymouth

de Boyer Montégut C, Madec G, Fischer AS, Lazar A, Iudicone D (2004) Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. J Geophys Res 109:C12003, doi:10.1029/2004JC002378

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27:2194–2200

Fonseca VG, Carvalho GR, Sung W, Johnson HF, and others (2010) Second-generation environmental sequencing unmasks marine metazoan biodiversity. Nat Commun 1:98

Hajibabaei M, Shokralla S, Zhou X, Singer GAC, Baird DJ (2011) Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. PLoS ONE 6:e17497

Hays GC, Richardson AJ, Robinson C (2005) Climate change and marine plankton. Trends Ecol Evol 20:337–344

Hirai J, Kuriyama M, Ichikawa T, Hidaka K, Tsuda A (2015) A metagenetic approach for revealing community structure of marine planktonic copepods. Mol Ecol Resour 15: 68–80

Huse SM, Welch DM, Morrison HG, Sogin ML (2010) Ironing out the wrinkles in the rare biosphere through improved OTU clustering. Environ Microbiol 12:1889–1898

Irigoien X, Huisman J, Harris RP (2004) Global biodiversity patterns of marine phytoplankton and zooplankton. Nature 429:863–867

Ji Y, Ashton L, Pedley SM, Edwards DP, and others (2013) Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. Ecol Lett 16:1245–1257

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772–780

Kidachi T, Itoh H (1979) Distribution and structure of macroplankton communities in the Kuroshio and coastal region, south of Honshu, during spring season. Bull Tokai Reg Fish Res Lab 97:1–119 (in Japanese)

Kunin V, Engelbrektson A, Ochman H, Hugenholtz P (2010) Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. Environ Microbiol 12:118–123

Landry MR, Al-Mutairi H, Selph KE, Christensen S, Nunnery S (2001) Seasonal patterns of mesozooplankton abundance and biomass at Station ALOHA. Deep-Sea Res II 48:2037–2061

Lindeque PK, Parry HE, Harmer RA, Somerfield PJ, Atkinson A (2013) Next generation sequencing reveals the hidden diversity of zooplankton assemblages. PLoS ONE 8:e81327

Longhurst AR (1985) The structure and evolution of plankton communities. Prog Oceanogr 15:1–35

Longhurst AR (2007) Ecological geography of the sea. Academic Press, London

Margulies M, Egholm M, Altman WE, Attiya S, and others (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437:376–380

Mauchline J (1998) The biology of calanoid copepods, Vol 33. Academic Press, San Diego, CA

McGowan JA (1971) Oceanic biogeography of the Pacific. In: Funnell BM, Riedel WR (eds) The micropaleontology of the oceans. Cambridge University Press, Cambridge, p 3–74

McGowan JA (1974) The nature of oceanic ecosystems. In: Miller CB (ed) The biology of the Pacific Ocean. Oregon State University Press, Corvallis, OR, p 9–28

McGowan JA, Walker PW (1979) Structure in the copepod community of the North Pacific central gyre. Ecol Monogr 49:195–226

Norton EL, Goetze E (2013) Equatorial dispersal barriers and limited population connectivity among oceans in a planktonic copepod. Limnol Oceanogr 58:1581–1596

Pearman JK, El-Sherbiny M, Lanzén A, Al-Aidaroos AM, Irigoien X (2014) Zooplankton diversity across three Red Sea reefs using pyrosequencing. Front Mar Sci 1:27

Richardson AJ, Walne AW, John AWG, Jones TD, and others (2006) Using continuous plankton recorder data. Prog Oceanogr 68:27–74

Roemmich D, McGowan J (1995) Climatic warming and the decline of zooplankton in the California current. Science 267:1324–1326

Rombouts I, Beaugrand G, Ibanez F, Chiba S, Legendre L (2009) Global latitudinal variations in marine copepod diversity and environmental factors. Proc R Soc B 276: 3053–3062

Rombouts I, Beaugrand G, Ibañez F, Gasparini S, Chiba S, Legendre L (2010) A multivariate approach to large-scale variation in marine planktonic copepod diversity and its environmental correlates. Limnol Oceanogr 55:2219–2229

Schloss PD, Westcott SL, Ryabin T, Hall JP, and others (2009) Introducing MOTHUR: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75:7537–7541

Schnack-Schiel SB, Mizdalski E, Cornils A (2010) Copepod abundance and species composition in the eastern subtropical/tropical Atlantic. Deep-Sea Res II 57:2064–2075

Shimode S, Takahashi K, Shimizu Y, Nonomura T, Tsuda A (2012) Distribution and life history of the planktonic copepod, *Eucalanus californicus*, in the northwestern Pacific: Mechanisms for population maintenance within a high primary production area. Prog Oceanogr 96:1–13

ter Braak CF, Verdonschot PM (1995) Canonical correspondence analysis and related multivariate methods in aquatic ecology. Aquat Sci 57:255–289

Terazaki M, Tomatsu C (1997) A vertical multiple opening

and closing plankton sampler. J Adv Mar Sci Technol Soc 3:127–132

Turner JT (2004) The importance of small planktonic copepods and their roles in pelagic marine food webs. Zool Stud 43:255–266

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 73:5261–5267

Welschmeyer NA (1994) Fluorometric analysis of chlorophyll *a* in the presence of chlorophyll *b* and pheopigments. Limnol Oceanogr 39:1985–1992

Williamson M, McGowan JA (2010) The copepod communities of the north and south Pacific central gyres and the form of species-abundance distributions. J Plankton Res 32:273–283

Woodd-Walker RS, Ward P, Clarke A (2002) Large-scale patterns in diversity and community structure of surface water copepods from the Atlantic Ocean. Mar Ecol Prog Ser 236:189–203

Editorial responsibility: Robert Condon, Wilmington, North Carolina, USA

Submitted: November 28, 2014; Accepted: June 24, 2015
Proofs received from author(s): August 14, 2015