

This is a peer-reviewed, author's accepted manuscript of the following research article: Xu, Z., Chen, J., Li, Y., Shekariz, E., Wu, W., Chen, B., & Liu, H. (2023). High microeukaryotic diversity in the cold-seep sediment. *Microbial Ecology*. <https://doi.org/10.1007/s00248-023-02212-y>

High microeukaryotic diversity in the cold-seep sediment

**Zhimeng Xu^{1,2}, Jiawei Chen^{1,2}, Yingdong Li², Erfan Shekariz², Wenxue Wu³, Bingzhang Chen^{1,4},
Hongbin Liu^{1,2,5,6*}**

¹Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou, China

²Department of Ocean Science, The Hong Kong University of Science and Technology, Hong Kong, China.

³State Key Laboratory of Marine Resource Utilization in South China Sea, Hainan University, Haikou, China.

⁴Department of Mathematics and Statistics, University of Strathclyde, Glasgow, UK

⁵Department of Ocean Science and Hong Kong Branch of the Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), The Hong Kong University of Science and Technology, Hong Kong, China.

⁶CAS-HKUST Sanya Joint Laboratory of Marine Science Research, Sanya, China.

*Corresponding author: Hongbin Liu (Email: liuhb@ust.hk)

Acknowledgements

We thank all members on the cruise HYDZ6-202102. We thank the crew and captain of R/V “Haiyangdizhi VI” for their great help during the work at the sea. We also thank Dr. Guangyuan Lu for his kind help on measuring environmental parameters.

Abstract

Microeukaryotic diversity, community structure and their regulating mechanisms remain largely unclear in chemosynthetic ecosystems. Here, using high-throughput sequencing data of 18S rRNA genes, we explored microeukaryotic communities from the Haima cold seep in the northern South China Sea. We compared three distinct habitats: active, less active and non-seep regions, with vertical layers (0–25 cm) from sediment cores. The results showed that seep regions harbored more abundant and diverse parasitic microeukaryotes (e.g., Apicomplexa and Syndiniales) as indicator species, compared to nearby non-seep region. Microeukaryotic community heterogeneity was larger between habitats than within habitat, and greatly increased when considering molecular phylogeny, suggesting the local diversification in cold-seep sediments. Microeukaryotic α -diversity at cold seeps was positively increased by metazoan richness and dispersal rate of microeukaryotes, while its β -diversity was promoted by heterogeneous selection mainly from metazoan communities (as potential hosts). Their combined effects led to the significant higher γ -diversity (i.e., total diversity in a region) at cold seeps than non-seep regions, suggesting cold-seep sediment as a hotspot for microeukaryotic diversity. Our study highlights the importance of microeukaryotic parasitism in cold-seep sediment and has implications for the roles of cold seep in maintaining and promoting marine biodiversity.

Keywords

Cold seep, Microeukaryotic diversity, Parasitism, Local diversification

1 Introduction

Unlike euphotic zones mainly characterized by photosynthesis, cold seeps are distinct seafloor ecosystems facilitated by chemoautotrophic microbes utilizing methane, sulfide and other hydrocarbons [1]. These biotic processes largely regulate marine biogeochemical cycles, especially for the flux of methane from the sediments into the water columns, emphasizing the importance of investigating the diversity, structure and functions of microbial communities at both sea floor and sediment of cold seeps [2, 3]. Cold seeps are shown as hotspots for not only megafauna but also microorganisms, including bacteria, protists and virus, by harboring diverse endemic species and acting as “island-like” habitats to maintain biodiversity and promote species diversification [4–6]. Shaped by local environments, microbial communities at cold seeps are patchily distributed and can vary greatly at spatial scales of only tens of meters [7, 8]. In the meantime, seafloor topology and ocean currents can influence the dispersal of seep species, finally shaping their genetic connectivity between habitats and determining their community spatial patterns [9]. Therefore, deciphering the mechanisms underlying the biogeography of cold seep microbial communities could have profound implications for organismal evolution and conservation of biodiversity in the deep sea.

Although invisible due to their small size, microeukaryotes (including protists and fungi), represent the vast bulk of eukaryotic diversity and biomass in most ecosystems including both water and sediment [10, 11]. They contain groups with different trophic modes, such as autotrophs (e.g., diatoms) as main primary producers in the pelagic ocean, heterotrophs (e.g., ciliates) as bacterial grazers, parasites (e.g., Apicomplexa) with wide range of hosts, osmotrophs (e.g., fungi) decomposing organic matters and mixotrophs (e.g., dinoflagellates) with the ability to switch between autotrophic and heterotrophic modes. This states their crucial and diverse roles in the ecosystems, such as transferring energy and materials in food webs (between microbial loop and metazoan food web) and global biogeochemical cycles [11–13]. Microeukaryotes, compared with bacteria and fauna, are much less studied at cold seep and have been sporadically reported with preliminary evidences showing that fungi and ciliates are likely to be dominant with great numbers of novel lineages [14–17]. One previous study showed that type of substrate, seep activity, sulfide concentration significantly affected microeukaryotic community diversity and compositions [18], while effects from other chemical factors (e.g., methane and nutrients) and processes

(e.g., dispersal) remain unclear.

Although cold seeps are known as hotspots for biodiversity, several studies reported that both bacterial and fauna richness in cold-seep sediment were lower than non-seep or less active seep area [7, 19], opposite to the findings of archaea and nematodes [20, 21]. This discrepancy could be mainly attributed to (1) the different mechanisms regulating different taxa (e.g., between bacteria and protists) even under the same environment [22–24]; (2) the differences in sampling strategies, because some studies only sampled surface layer of sediment while others included deeper layers such as the sulfide-methane transition zone which is of great environmental gradients and biotic active layer in cold-seep sediment [7, 18, 20, 25]. Thus, we suggest that comparing diversity under sufficient sampling efforts and disentangling the fundamental processes underlying the communities are necessary to know the roles of cold-seep sediment on shaping microeukaryotic diversity.

It is widely acknowledged that both deterministic and stochastic processes play fundamental roles in microbial community assembly, determining the diversity and coexistence of species [26–29]. Deterministic processes are niche-based mechanisms and include environmental filtering and interspecific interactions (e.g., competition, facilitation, mutualism and predation), while stochastic processes are neutral-based and refer to random changes in the relative abundance of species, involving random birth, death and dispersal events [30–32]. According to the species-area relations [33], alpha (α , species richness in a local community) and beta (β , compositional turnover between communities) diversity can be mainly influenced by several effects. For example, improved environmental conditions can favor higher α -diversity (the habitat quality effect), while dispersal at low or intermediate rates can positively increase α -diversity but have negative effects at high rates [33, 34]. Microbial community β -diversity has been frequently studied by comparison between environmental selection (represented by measured environmental factors) and dispersal related processes (e.g., dispersal limitation, which can be reflected by distance decay relationships, DDR) [35–37]. From the niche-inferred aspect, selection acts on β -diversity from two opposite directions depending on environmental heterogeneity: heterogeneous selection increases β -diversity while homogeneous selection decreases β -diversity; From the neutral aspect, dispersal limitation (dispersal at low rates) increases β -diversity while homogenizing dispersal (dispersal at high rates) decreases β -diversity [29, 30]. Still, a central issue in microbial ecology is to

quantify the relative importance of these processes in regulating community diversity, composition and biogeography [26, 28, 38], which hasn't been well explored in microeukaryotes in cold-seep sediments.

The aim of this study was to investigate the microeukaryotic communities in cold-seep sediment, with the test of whether cold seeps could be hotspots of microeukaryotic diversity. We recovered the microeukaryotic communities using DNA signatures collected from the sediment of the Haima cold seep located in the northern slope of the South China Sea and currently active [25, 39]. We compared the microeukaryotic communities among three distinct habitats, representing active, less active and non-seep regions respectively. Given that microbial communities in cold-seep sediment are greatly heterogeneous across spaces even at small distances [7], we argue that α -diversity alone cannot represent the average or total species richness at one habitat especially with limited sampling efforts. Instead, estimating the γ -diversity (total species richness in a region) can provide more convincing and direct answers to our proposed test. We generated a framework to compare and decipher the γ -diversity between cold seep and non-seep regions, which was mainly based on two theoretical points: Firstly, γ -diversity is decided by pure sampling efforts and the combined effects of α - and β -diversity [33, 40]. Secondly, as introduced above, α -diversity can be enhanced by habitat quality or low/intermediate dispersal rate, while β -diversity is mainly controlled by selection and dispersal processes. By analyzing the community data and following this framework, we would like to clarify: (1) do cold-seep sediment have unique microeukaryotic communities with indicator groups or species in functional or trophic aspects; (2) whether microeukaryotic communities are largely different within and between habitats, as described in bacterial communities [7]; (3) does cold-seep sediment host higher γ -diversity than nearby? If so, what are the main factors making cold-seep sediment as a hotspot of microeukaryotic diversity.

2 Methods

2.1 Sample collection and pore water measurement

Samples were collected from the Haima cold seep (16.72°N, 110.46°E, depth of ~ 1400 m) at the northern South China Sea during the cruise HYDZ6-202102 on R/V "Haiyangdizhi VI" in May 2021 (Fig. 1). Sediment push cores were retrieved using the remotely operated underwater vehicle (ROV) at 3 distinct habitats: ROV1, mussel bed, characterized as massive mussels and arthropods (e.g., crabs and shrimps) and visible bubbling of seepage, showing high activity at seep region; ROV3, clam bed, with patchy

distribution of *Archivesica* spp. (Bivalvia: Vesicomidae) at seep region [25], treated as a less active seep region; ROV5, nearby non-seep region (10 – 20 km from ROV1 and ROV3) with flat seafloor and few megafauna. For each push core (9 cores in ROV1, 6 cores in ROV3 and 5 cores in ROV5), samples were vertically taken from surface (i.e., 0 cm) to 25 cm depth at every 5 cm, ideally generating 6 local communities. All samples were frozen immediately after taken by ROV and stored at -80°C until further use.

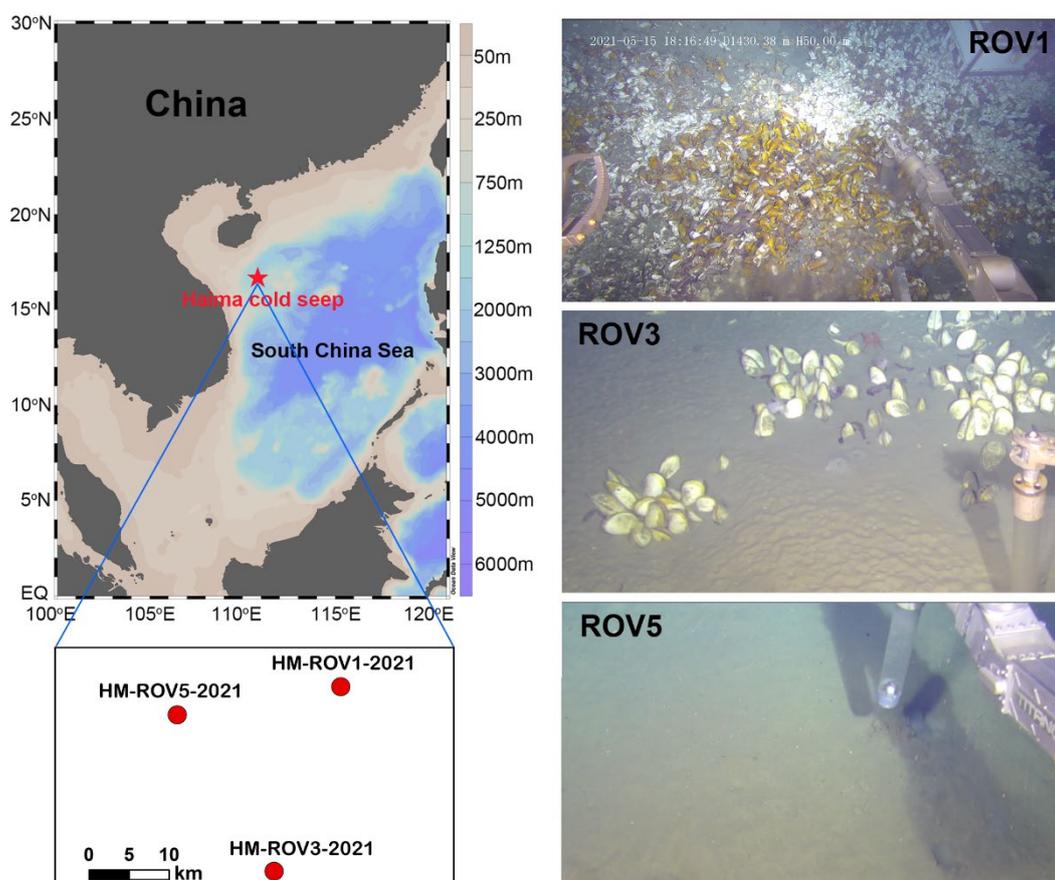


Figure 1 | Location of sampling sites and characters of habitats. Samples were collected from Haima cold-seep sediment in the northwest South China Sea. Three habitats were sampled: ROV1 (highly active seep region with massive mussels and crabs), ROV3 (less active seep region with patch distributions of clams) and ROV5 (non-seep region with few megafauna).

Pore water for each sample was extracted from the sediment cores using Rhizon moisture samplers [73] and measured for the following chemical factors in the lab of the Third Institute of Oceanography,

Ministry of Natural Resources in China. Concentration of methane (CH₄) was measured by gas chromatography Agilent 6850 (Agilent Technologies, USA). Sulfide (S²⁻) and phosphate (PO₄³⁻) were measured using a discrete auto-analyzer Smartchem200 (Alliance, France). Anions and cations, i.e., sulfate (SO₄²⁻) and ammonium (NH₄⁺), were analyzed by an ion chromatography ISC-1000 (Thermo scientific, USA). Concentration of dissolved inorganic carbon (DIC) and isotopic analysis of ¹³C of DIC were tested by a mass spectrometer Delta V Advantage (Thermo scientific, USA) with PoraPlotQ column (Agilent Technologies, USA). The ¹³C-DIC values are reported in the conventional δ notation in permil (‰) relative to V-PDB (Vienna Pee Dee Belemnite). For each habitat, an additional long push core with 70 cm length was used to show the vertical distributions of the above environmental factors.

2.2 DNA extraction, PCR, sequencing and reads processing

DNA was extracted from each sediment sample using DNeasy PowerSoil Pro Kit (QIAGEN, USA) according to the manufacturer's protocol. Quality and concentration of extracted DNA were measured by Qubit dsDNA Assay Kit in Qubit 2.0 Fluorometer (Life Technologies, USA) and checked by 1% agarose gel electrophoresis. Polymerase chain reaction (PCR) was performed with barcoded universal primers targeting the hypervariable V4 region of the 18S rRNA gene: 528F (5'-GCGGTAATTCCAGCTCCAA) and 706R (5'-AATCCRAGAATTTACCTCT) [74]. PCR mixtures (25 µl) were prepared in triplicate, each containing 2.5 µl of 10 PCR buffer, 0.75 µl of 10 mM MgCl₂, 0.5 µl of 10 mM dNTP mix, 0.5 µl of each primer (10 µM) and 1 U of Invitrogen Platinum Taq DNA polymerase (Life Technologies, USA). The PCR reaction condition was set as: an initial denaturation of 3 min at 94°C, followed by 30 cycles of 30 s at 94°C, 30 s at 60°C, 1 min at 72°C and a final cycle of 5 min at 72°C. PCR products were pooled together into library and sequenced by a HiSeq 2500 System (Illumina, USA) with 2×250 bp paired end read configurations.

Raw reads were processed using QIIME2 (<https://docs.qiime2.org/>) with tutorial pipelines [75]. In brief, after removal of barcode and primer, sequences (with paired ends for each sample) were imported into QIIME2 (v 2019.10), visually inspected for sequence quality after demultiplexing, and processed with DADA2 to remove contamination, trim reads, correct errors, merge read pairs and remove chimeras. Representative ASV (amplicon sequence variants) sequences and their abundances were extracted by feature-table (file "table.qza"). Phylogenetic trees, both rooted and unrooted, were built with

representative ASVs based on their nucleic acid sequences, for further analysis when phylogenetic distances between ASVs are needed. A naive Bayes classifier was fitted with 18S rDNA sequences extracted from SILVA v132 database [76] according to the PCR primers used. Then, the representative ASV sequences were classified with detail taxonomy information using the trained classifier. ASVs with taxonomic assignment of prokaryotes, metazoan and land plants were removed (metazoan ASVs will be used for parts of the downstream analysis separately).

2.3 Community composition analysis

R (v. 4.1.1) software with different packages were used in the following analyses [77]. Relative abundance of microeukaryotic sequences at the phylum or class level was calculated to show the abundant groups for samples at each habitat. At the ASV level, community similarity between two samples were calculated using Bray-Curtis distance (“vegdist” function in “vegan” package) and visualized on a non-metric multi-dimensional scaling (NMDS) map using the first two dimensions (“monoMDS” function in “vegan” package). Effects from two factors, depth and habitat type, on community similarity (Bray-Curtis) were tested, respectively, by analysis of similarity (ANOSIM) (“anosim” function in “vegan” package, with permutations = 999) [78]. Tests were performed for both individual and combined habitats, and a larger ANOSIM-R value suggested greater difference among groups of samples divided by depth or habitat.

To show the representative functional or trophic groups of each habitat, we firstly draw a ternary plot based on the relative abundance of ASVs at each habitat (“ggtern” function in “ggtern” package) [79]. These ASVs, according to their taxonomy, were classified by the most potential trophic modes as reported in previous studies [80, 81]. Secondly, as a support, we performed indicator species analysis (ISA) to find the ASVs that could be indicator species of each habitat, using the “multipatt” function (func = “IndVal”, nperm = 9999) in “indicspecies” package [82]. ASVs with *p*-value less than 0.05 were treated as indicator species, with higher IndVal value suggesting more strongly associated to one habitat.

2.4 Diversity estimation and determining factors

Here we used a framework combining α - and β - and γ -diversity to compare the microeukaryotic communities between habitats. We defined α -diversity as the species number of a local community

revealed by each sample (i.e., a certain layer of a sediment core) and it was measured by ASV richness (i.e., observed ASV number), Chao1 (used as predicted richness in situ) and phylogenetic diversity (PD, which considers phylogenetic distance between ASVs) [48, 83]. In brief, PD is a measure of the evolutionary relationship between species and provides a metric of biodiversity that accounts for evolutionary distances between co-occurring species [84]. Phylogenetic closed species are supposed to have similar evolutionary potential, while more distantly related species differ more in their potential.

β -diversity referred to the community dissimilarity, calculated by “vegdist” function (“vegan” package) using Bray-Curtis, Jaccard and UniFrac distances respectively, between any two local communities at a habitat [78, 85]. In particular, community differences among habitats were compared between using Bray-Curtis and UniFrac distances (abundance unweighted), to showing the importance of phylogenetic distance between species in studying the community heterogeneity at cold seep regions [86]. γ -diversity here represented the total species richness at one habitat, and it was calculated by the total observed ASV richness from a whole ROV. To avoid pure sampling effects on comparing γ -diversity (due to the unequal number of samples in the three ROVs), we plotted an ASV accumulation curve for each habitat (“specaccum” function in package “vegan”, method = “random”) [87]. The ASV accumulation curve displayed the number of local communities (randomly sampled) on the x-axis and observed ASV richness on the y-axis, allowing us to compare cumulative diversity with equal numbers of samples between habitats.

For factors controlling α -diversity, we firstly evaluated the influences from both abiotic and biotic aspects. Significant correlations between measured abiotic environmental factors and microeukaryotic α -diversity indices (i.e., richness, Chao1 and PD) were indicated by Spearman correlation p -values ($p < 0.05$). Both linear model and generalized linear model (GLM, by “mgcv” package in R) were used to evaluate the effects of bacterial (16S rDNA data from the same samples, unpublished) and metazoan richness on microeukaryotic richness [88]. Then, we analyzed the effects of species dispersal on α -diversity. Distance decay relationships (DDR) between geographic distance and community similarity, which could represent dispersal limitation, was tested at each habitat. Here we only tested the DDR vertically within each sediment core (due to the unknown horizontal distance between two sediment cores). Further, treating each sediment core as a metacommunity, we estimated the dispersal rate (m) following the

neutral community model (NCM) [89, 90]:

$$Freq_i = 1 - I(1/N | N * m * p_i, N * m * (1 - p_i))$$

Where $Freq_i$ is the occurrence frequency of ASV i across metacommunity; N is the number of individuals per community; m is the estimated immigration rate; p_i is the mean relative abundance of ASV i across metacommunity; $I()$ is the probability density function of beta distribution. The NCM is based on predicting the relationship between the frequency of ASVs (i.e., site occupation frequency in the metacommunity) and their abundances across the metacommunity [91]. It predicts that abundant taxa are more likely to be dispersed by chance and widespread (i.e., high frequency) in the metacommunity, while rare taxa would be lost due to drift. In the model, a high immigration rate (m) indicates the high species dispersal in the metacommunity (i.e., low dispersal limitation). We also calculated the dispersal rate of an entire habitat (as one metacommunity) with all samples from it. Dispersal rates, from both sedimental core and habitat levels, were compared among the three habitats. Correlation between ASV richness (on average in a sediment core) and dispersal rate was tested by linear regression.

We used null model analysis [26, 38] to estimate the relative contribution of ecological processes (especially heterogeneous selection) on regulating the microeukaryotic community β -diversity. In brief, phylogenetic turnover using abundance weighed β -mean nearest taxon distance (β MNTD) metric was measured, which quantifies the mean phylogenetic distances between two evolutionary-closest ASVs in two communities:

$$\beta\text{MNTD} = 0.5 \left[\sum_{i_k=1}^{n_k} f_{i_k} \min(\Delta_{i_k j_m}) + \sum_{i_m=1}^{n_m} f_{i_m} \min(\Delta_{i_m j_k}) \right]$$

where f_{i_k} is the relative abundance of ASV i in community k , n_k is the number of ASVs in k and $\min(\Delta_{i_k j_m})$ is the minimum phylogenetic distance between ASV i from community k and all ASVs j from community m . β MNTD was calculated using the function “comdistnt” (with abundance weighted) in package “picante” [92]. β MNTD values higher than expected by chance indicate heterogeneous (or variable) selection in community assembly while β MNTD values lower than expected by chance indicate homogeneous selection. The null model expectation was performed using 999 randomizations, and deviation between the observed β MNTD and the mean of null model distribution is shown as β -nearest taxon index (β NTI). A significant deviation (i.e., $|\beta\text{NTI}| > 2$) indicates the dominance of selection processes: $\beta\text{NTI} < -2$ indicates significantly less phylogenetic turnover than expected (i.e., homogeneous

selection) while $\beta\text{NTI} > 2$ indicates significantly more phylogenetic turnover than expected (i.e., heterogeneous selection). If the deviation is low (i.e., $|\beta\text{NTI}| < 2$), it means the β -diversity of communities could be structured by random processes such as dispersal and drift [28, 38]. To know the factors mainly contributed to heterogeneous selection and regulated the microeukaryotic communities in cold-seep sediment, we analyzed the correlations between microeukaryotic β -diversity (Bray-Curtis dissimilarity) and differences in abiotic factors (measured chemical parameters) and biotic factors (β -diversity of metazoan and bacterial communities) (by linear regression models).

3 Results

3.1 Vertical distributions of chemical factors

Microeukaryotic communities in our study were collected from samples of sediment with depth ranging from 0 to 25 cm. We used the chemical concentrations (i.e., $\delta^{13}\text{C}$ of DIC, CH_4 , SO_4^{2-} , S^{2-} , NH_4^+ and PO_4^{3-}) measured from a long sediment push core (0 – 70 cm) to provide a clearer pattern of the vertical chemical properties at each habitat (Fig. S1). In brief, $\delta^{13}\text{C}$ of DIC was the lowest at ROV1 (-34.33‰, on average) with a minimum point at ~ 5 cm depth (~ -45‰). Concentration of CH_4 was consistently the highest at ROV1 (10613±14477 mg/kg) compared with ROV3 (384±62 mg/kg) and ROV5 (240±19 mg/kg), with a generally increasing pattern from surface to 25 cm depth. In contrast, concentration of SO_4^{2-} was the lowest at ROV1, decreasing drastically from 2096.32 mg/L at surface to 70.30 mg/L at 10 cm depth. S^{2-} concentration was the highest at ROV1 (289±266 mg/L) with an increasing trend from surface to 10 cm depth. No clearly comparable patterns were observed for the concentrations of NH_4^+ and PO_4^{3-} among the three ROVs. Values from other sediment cores (0 – 25 cm) followed the similar pattern (Table S1).

3.2 Sequencing results, community composition and indicator species

After sequencing, quality control and merging of paired reads, 6,568,903 sequences of 18S rDNA V4 region were obtained. By clustering sequences to ASVs, a total of 15393 ASVs were recovered from all samples (126 samples, ROV1: 54, ROV3: 43, ROV5: 29) with an average of 56628 sequences in each sample. After removal of metazoan (but used for parts of downstream analysis as potential hosts), terrestrial plants (e.g., Embryophyta) and unknown (cannot matched to any taxonomic level) ASVs, microeukaryotic communities were constructed by 126 samples with 6424 ASVs and 141,216 sequences (average of 30,642 sequences in each sample) (Table S2).

At a raw group level (phylum or class), Fungi (belonging to the super group Opisthokonta) mostly dominated the microeukaryotic communities at all ROVs with sequences averagely contributing to 44.65% at ROV1, 53.07% at ROV3 and 60.87% at ROV5 (Fig. S2 and Fig. S3a). Apicomplexa (Alveolata) sequences were more enriched at seep area, averagely contributing to 13.11% at ROV1, 9.63% at ROV3 and 3.80% at ROV5, while Retaria sequences (Rhizaria) was more abundant at ROV5 (5.58%) than ROV1 (2.04%) and ROV3 (1.39%). Proportion of ASV number of each group microeukaryotes showed the similar pattern to the sequence abundance, especially for the three groups (Fungi, Apicomplexa and Retaria) described above. ROV1 harbored the highest ASV number of 3529 (ROV3: 2832, ROV5: 1652) with the largest proportion of endemic (ROV-specific) ASVs of 69.06% (ROV3: 60.7%, ROV5: 58.29%) (Fig. S3b).

Parasitic species within Apicomplexa and Syndiniales were more enriched at seep area especially ROV1 while Retaria (phagotrophic) and some Fungi (osmotrophic) species were more abundant at non-seep area ROV5 (Fig. 2), in accord to the indicator species analysis of each habitat (Table S3–S6). We identified 29 indicator species of active cold-seep sediment (ROV1) ($p < 0.05$). For instance, ASV3771 (belonging to Syndiniales Group 1), ASV4397 (belonging to Novel Apicomplexa Class 1) had much high indicator values (IV, 0.59 and 0.54, respectively) and could be strong indicator species of ROV1. Notably, while indicator species of ROV1 were from diverse groups of microeukaryotes (e.g., Syndiniales, Chlorophyta, Apicomplexa, Dinoflagellata and Breviatea), indicator species of ROV5 were mostly affiliated to Fungi and Retaria.

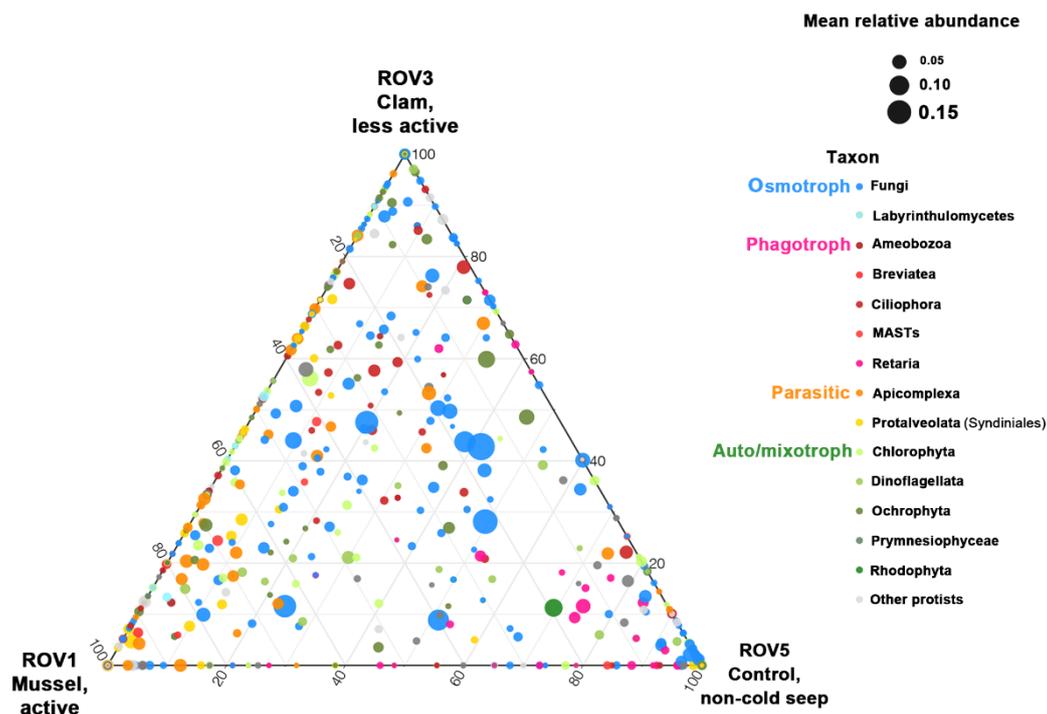


Figure 2 | Ternary plot of taxa in three ROVs. Each dot represents an ASV with abundance-weighted distribution in the three ROVs. Dots were colored according to corresponding taxon (class or phylum).

Size of dots was scaled by mean relative abundance of ASVs across all ROVs.

At the ASV level, community similarity of microeukaryotes was calculated using Bray-Curtis distance and visualized on a NMDS map to show the effects of both depth and habitat type on community composition patterns (Fig. S4a). Tested by ANOSIM, our results showed that, generally, habitat type had larger (ANOSIM-R) and more significant (ANOSIM-*p*) effects on microeukaryotic community structure, compared with depth (Fig. S4b).

3.3 Comparison of α -, β - and γ -diversity of microeukaryotes

Higher α -diversity of microeukaryotes was detected in cold-seep sediment (ROV1&3) than non-seep area (ROV5) (Fig. 3a). ASV richness of microeukaryotes at ROV1 (on average of 121) and ROV3 (on average of 114) were not significantly different between each other ($w = 1227$, $p = 0.63$, by Wilcoxon rank sum test as richness values were not normal distributions), but significantly higher than ROV5 (on average of 81) (both $p < 0.01$, Wilcoxon test). Chao1 index showed the same pattern (ROV1: 153, ROV3: 157, ROV5: 112, on average). However, phylogenetic diversity of microeukaryotes at ROV1 was

significantly higher than ROV3 (on average of 18.77), followed by ROV5 (on average of 15.98) (for all comparisons, $p < 0.05$, Wilcoxon test).

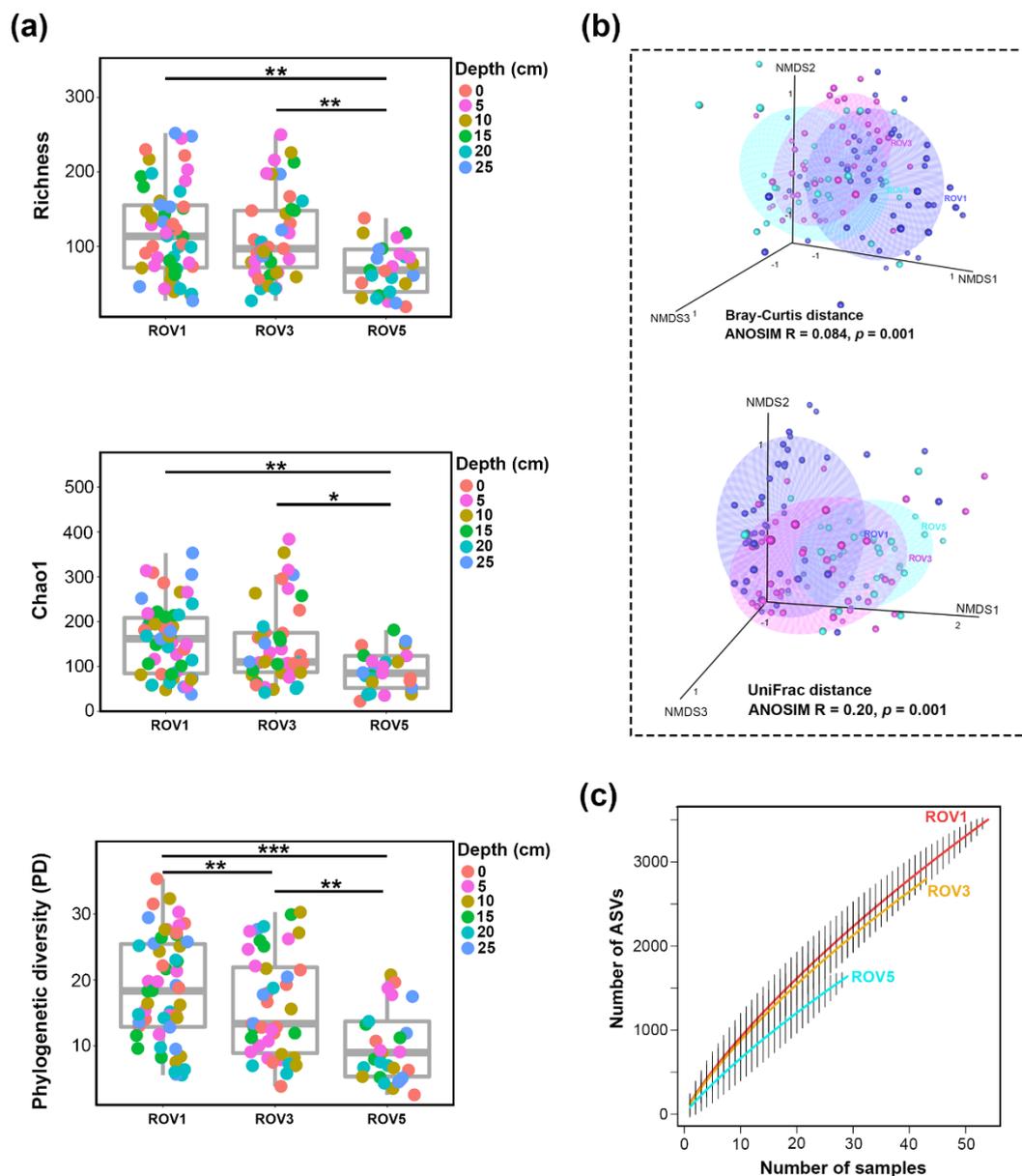


Figure 3 | Comparisons of α - and β -diversity between ROVs. Three indices of α -diversity (richness, Chao1 and PD) of microeukaryotes were compared between ROVs. PD: phylogenetic diversity (a); comparison of community difference between ROVs using Bray-Curtis and UniFrac distance, respectively. Community difference between ROVs was tested by ANOSIM, and a higher R value indicates a larger group-level difference (c); comparison of regional diversity (i.e., γ -diversity) between ROVs with randomly equal number of samples.

High microeukaryotic diversity in the cold-seep sediment

Microeukaryotic β -diversity (represented by Bray-Curtis dissimilarity) was the highest at ROV1 with the largest community dissimilarity (on average of 92.41%), followed by ROV5 (on average of 90.66%) and ROV3 (on average of 88.69%) (for all comparisons, $p < 0.01$, Wilcoxon test) (Fig. S5a). Community dissimilarity using Jaccard distance showed the similar pattern (ROV1 had the highest β -diversity, but no significant difference was found between ROV3 and ROV5) (Fig. S5b). Communities could be grouped by ROVs based on community similarity calculated by either Bray-Curtis or UniFrac (abundance unweighted) distance (ANOSIM test, $p < 0.05$). However, community difference between ROVs using UniFrac distance was much larger (ANOSIM $R = 0.20$) than using Bray-Curtis distance (ANOSIM $R = 0.084$), indicating the importance of local (within habitat) phylogenetic divergence on community β -diversity (Fig. 3b).

Microeukaryotic γ -diversity was estimated by the total observed number of ASVs (i.e., richness) from all detected samples. ROV1 showed the consistently highest γ -diversity with randomly equal number of samples between ROVs (Fig. 3c). After rarefied to the same sequence depth, rarefaction curve for each individual sample within a certain ROV was nearly saturated, suggesting the enough sequencing efforts for all samples (Fig. S6a-c). ROV1 was also found with the highest γ -diversity at ROV level, followed by ROV3 and ROV5, by comparisons of ASV richness from rarefying communities at the same sequence depth ($n = 32510$) (Fig. S6d).

3.4 Factors controlling the microeukaryotic diversity

We firstly found that there was no significant correlation between measured chemical factors ($\delta^{13}\text{C}$ of DIC, CH_4 , SO_4^{2-} , S^{2-} , NH_4^+ and PO_4^{3-}) and α -diversity (observed richness, Chao1 and phylogenetic diversity PD) of microeukaryotes (Spearman correlation $p > 0.05$), suggesting that the effects from abiotic habitat quality on α -diversity could be excluded here (Fig. 4a). Instead, we found strong effects from biotic factors on shaping microeukaryotic richness, especially from metazoan richness which explained 58.1% of the variation (indicated by generalized additive model, GAM) (Fig. 4b).

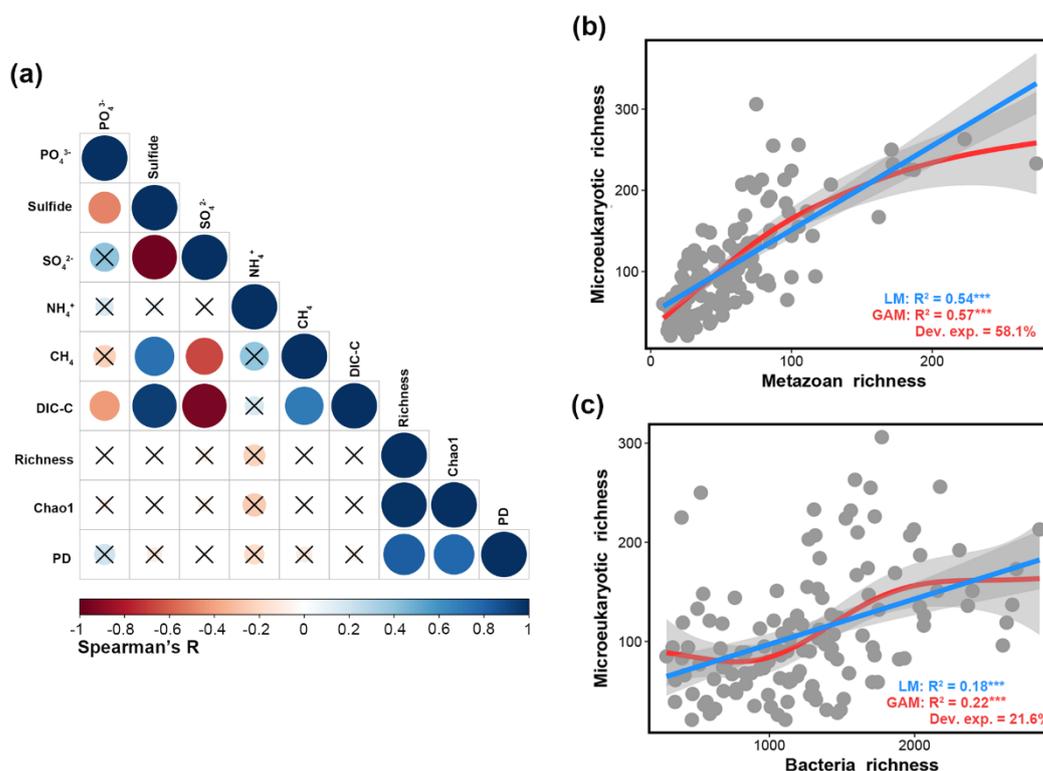


Figure 4 | Effects of abiotic and biotic factors on microeukaryotic α -diversity. Effects of measured abiotic environmental factors on microeukaryotic α -diversity (richness, Chao1 and PD) were tested by Spearman correlation coefficient. Spearman's R close to 1 refers to strong positive correlation while R close to -1 indicates strong negative correlation. Significant correlation ($p < 0.05$) was marked with filled circles (larger size represents smaller p -value) while a cross mark indicates not significant **(a)**. Linear model (blue line) and generalized additive model (GAM) (red line) were used to estimate the effects of biotic (metazoan and bacteria richness) factors on microeukaryotic richness. For GAM, R^2 is the adjusted value. Dev. exp.: deviance explained **(b, c)**.

Then, we showed that dispersal rate of species was higher at cold seep area than non-seep area (Fig. 5). Vertically from sediment cores (0 cm to 25 cm, including 6 samples), we found that microeukaryotic community similarity decreased significantly with increasing depth difference only at ROV5, showing the distance-decay relations (DDR), suggesting the roles of dispersal limitation at non-seep area (Fig. 5a). Dispersal rate of species at both sediment core level (ROV1: 0.065, ROV3: 0.076, ROV5: 0.023, on average, respectively) and ROV (habitat) level (ROV1: 0.018, ROV3: 0.015, ROV5: 0.0081) were significantly higher at cold seep area than non-seep area (Welch two sample t-test, $p < 0.05$) (Fig. 5b). With samples from all ROVs, we found that there was a positive correlation between dispersal rate of

species and average richness of microeukaryotes at a sediment core (Linear regression $R^2 = 0.58$, $p < 0.001$) (Fig. 5c).

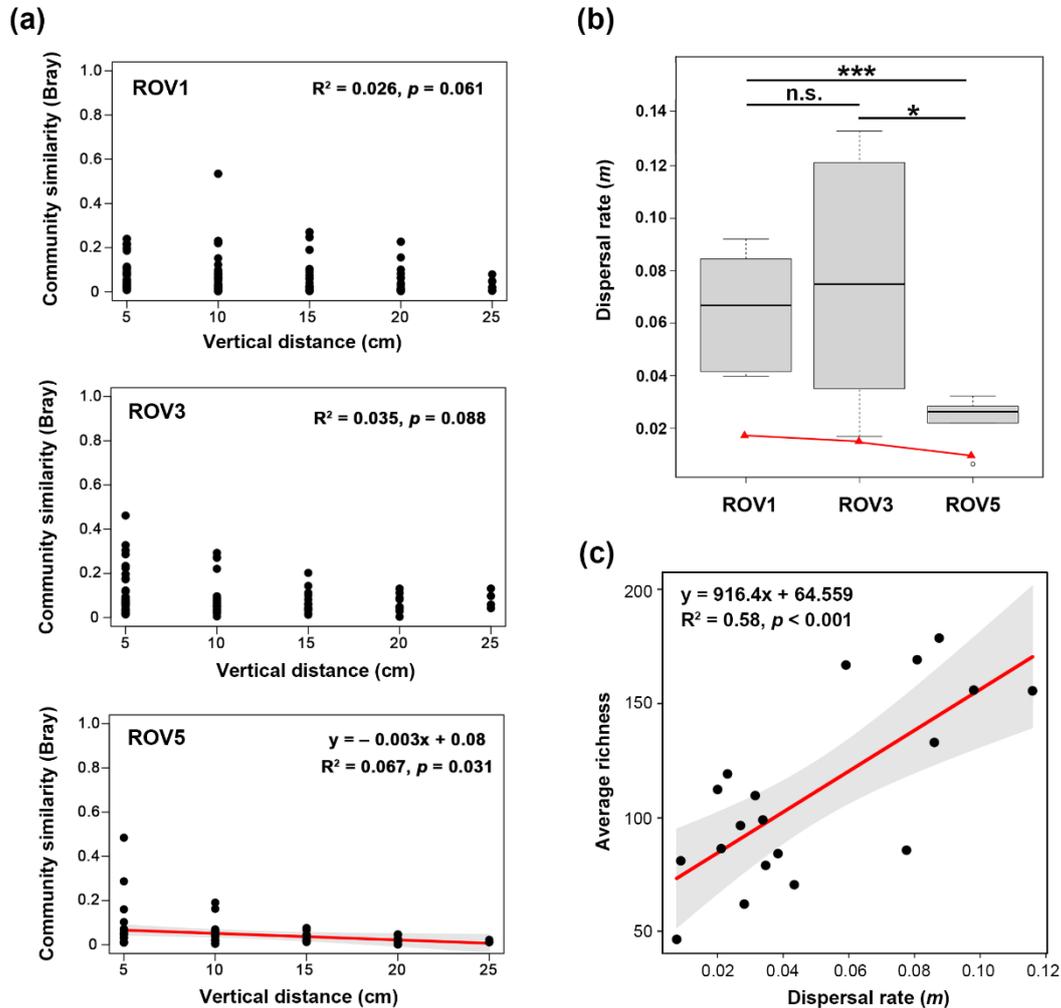


Figure 5 | Estimation of dispersal and relations between dispersal rate and microeukaryotic richness. Distance-decay relationship of community similarity (DDR) and geographic distance was used to reflect dispersal limitation **(a)**, comparisons of dispersal rate from both sediment core (boxplot) and ROV (red line) scales **(b)** and correlations between dispersal rate and average microeukaryotes richness in each sediment core **(c)**.

Using phylogenetic turnover (i.e., β NTI) based null model, we found that heterogeneous selection contributed more to the microeukaryotic community assembly processes at ROV1 (16%) than ROV3 (8%) and ROV5 (9%), which could explain the highest β -diversity at ROV1 (Fig. 6). Further, difference

of metazoan communities, compared to abiotic factors and bacterial communities, had the strongest positive correlation to the β -diversity of microeukaryotic communities at ROV1 (linear regression, $R^2 = 0.11$ and 0.21 for metazoan abundance weighted and unweighted, respectively) (Fig. S7), suggesting that the relatively higher heterogeneous selection at ROV1 might be attributed to the heterogeneity of metazoan communities as potential hosts of microeukaryotes.

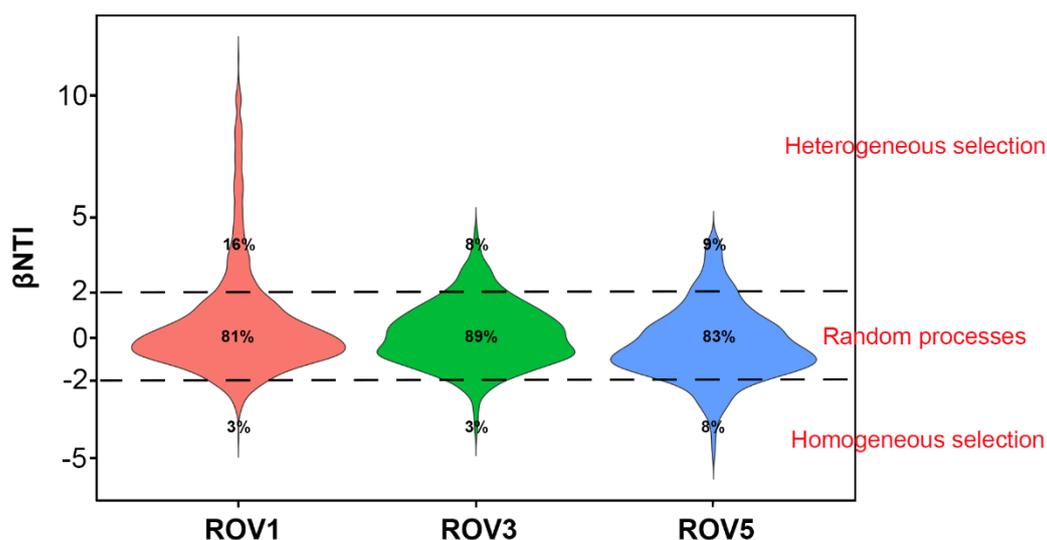


Figure 6 | Phylogenetic turnover (βNTI) in microeukaryotic communities calculated by null model. Different ecological processes were indicated by values of βNTI with $\beta\text{NTI} > 2$ indicating heterogeneous selection.

We further analyzed the metazoan communities. At the class level, sequences of Pteriomorpha (Mollusca, Bivalvia), Palpata, Scolecida and Oligochaeta (Annelida), and Enteropneusta (Hemichordata) were more enriched at seep sediment (especially ROV1) while Neoptera (Arthropoda, Insecta) was more abundant at non-seep sediment (ROV5) (Fig. S8a). At the ASV level, metazoan community differences between habitats were much higher in abundance unweighted results (i.e., using presence/absence data) than in weighted ones, which was indicated by higher ANOSIM-R ($0.14, p < 0.001$) of the former (Fig. S8b)

As a synthesis, we used a conceptual figure to show how factors controlled the α - and β -diversity which finally determined the γ -diversity of microeukaryotes at cold seep area (Fig. 7). Briefly, α -diversity was defined by species richness of a local community in each layer of a sediment core and β -diversity was

compared between communities within or between sediment cores, while γ -diversity was estimated as the total richness of species at a habitat (ROV). We listed several potential factors controlling microeukaryotic α -, β - and γ -diversity, respectively (Fig. 7a). After exclusion of the effects from abiotic habitat quality (on α -diversity), dispersal limitation (on β -diversity) and pure sampling effort (on γ -diversity), we attributed the highest γ -diversity at active cold-seep sediment to the combination of higher α -diversity increased by metazoan richness and dispersal rate (at low or intermediate level) and larger β -diversity caused by heterogeneous selection (Fig. 7b).

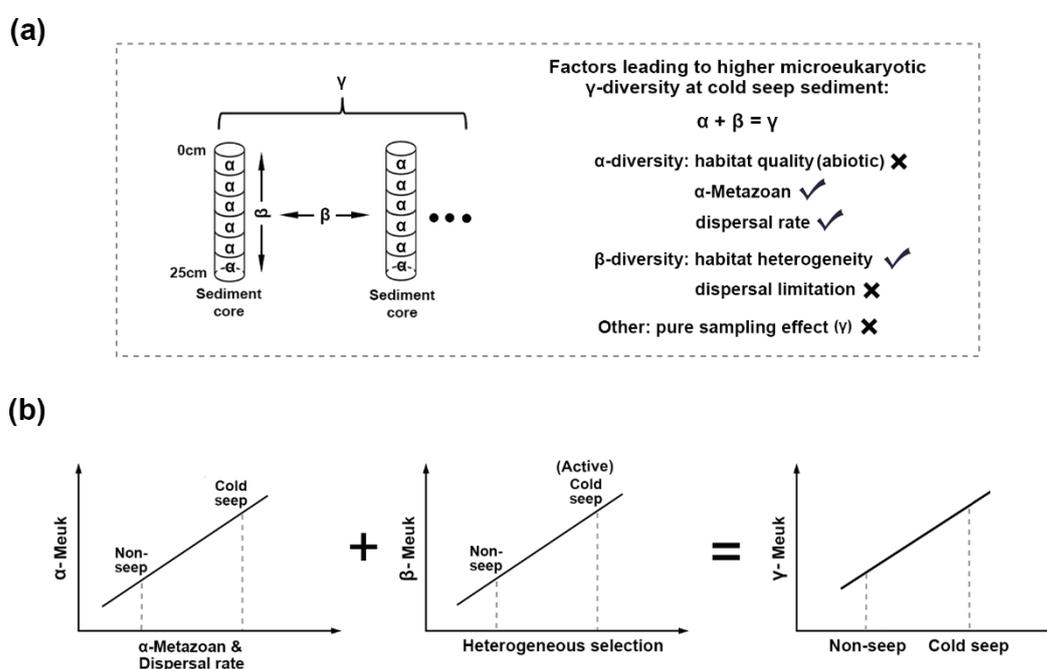


Figure 7 | A schematic figure showing the mechanisms shaping the highest microeukaryotic diversity in cold-seep sediment. Here, α -diversity is defined as ASV richness of a local community in a sediment core and β -diversity is the dissimilarity between communities within or between sediment cores, while γ -diversity is the total richness of species at one habitat (ROV). Several potential main factors controlling microeukaryotic α -, β - and γ -diversity are tested. Effects from habitat quality (on α -diversity), dispersal limitation (on β -diversity) and pure sampling effort (on γ -diversity) are excluded from our analyses (a). We finally attribute the highest γ -diversity at active cold-seep sediment to the combination of higher α -diversity increased by dispersal rate (at low or intermediate level) and larger β -diversity caused by heterogeneous selection (b).

4 Discussion

4.1 Microeukaryotic communities in cold-seep sediment are locally divergent with enhanced parasitism

Our study provides a comprehensive investigation on vertical and horizontal community structures of microbial eukaryotes in cold-seep sediment, with a comparison to nearby non-seep regions. Firstly, focused on the upper layers (0–25 cm) which are the most biochemically active depths, we found that fungi dominated in most samples at both seep and non-seep sediment. This is in accord to previous findings showing that basidiomycetous fungus dominated the microeukaryotic communities at methane cold seeps of both Sagami Bay and Kuroshima Knoll in the southern Ryukyu Arc [15, 16]. Together with several other groups, e.g., Labyrinthulomycetes, fungi were reported as the main osmotrophic-saprotrophic microeukaryotes and important for detrital decomposition not only in terrestrial soils but also in marine sediment [12]. They can catalyze larger organic matter and polymers, such as from detritus produced by megafauna and marine snow, into dissolved forms with extracellular enzymes, providing a redistribution loop of carbon recycling and nutrient shortcut [41]. Our results showed that while some fungi (ASVs) were widely distributed with high abundance across habitats, some other fungi were habitat-specific and could also be the local indicator species. We acknowledge that a further study of the variations of fungi communities between habitats together with phylogenetic analysis may enhance our understanding or provide new findings of the roles of fungi on biogeochemical cycles in cold-seep sediment, however, it is not within our main scopes of the present study.

Secondly, we found that parasitic groups, such as Apicomplexa and Syndiniales, were enriched at cold seep area relative to non-seep area, from both sequence abundance and ASV richness aspects. Parasitism dominated the protistan communities in both rainforest soil and hydrothermal vents with high diversity, but few is known about their compositions and roles in the cold-seep sediment [42–44]. Both Syndiniales and Apicomplexa are well known for their parasitic species with a large range of host including both invertebrates and vertebrates, playing important roles in controlling the population density of host animals [43, 45]. The indicator Apicomplexa species of ROV1 were closest to a novel class of Apicomplexa, together with increasing novel species of Apicomplexa being reported from hosts living at cold seep, it suggests that cold seep can serve as an ideal place for the diversification of parasitic

lineages [16, 18, 46, 47].

Our results showed that, when considering phylogeny, diversity among habitats were more different. From the α -diversity aspect, a community with high PD would have a high chance of containing a subset of species with great potential of evolution and being robust to future environmental changes [48, 49]. A previous global study of archaea showed that hydrothermal vents had higher PD than other habitats (e.g., soil and marine sediment) and the largest number of indicator species located close to the root of phylogenetic tree, serving as the largest reservoirs of archaeal diversity [50]. Similarly, in our study, highest phylogenetic diversity (PD) of microeukaryotes was found at active cold-seep sediment compared to less active seep and non-seep areas (Fig. 3a), suggesting that the active cold-seep sediment could be a stable habitat to hold diverse species under extreme environmental conditions. From the β -diversity aspect, we showed that microeukaryotic communities in the cold-seep sediment were locally divergent, which were reflected by the significant difference of community similarity among habitats. There result was same to the findings of previous studies reporting that bacterial communities had high dissimilarity both within and between different habitats (even much higher between habitats), suggesting that microbial communities at each unique seep region could be called as a microbiome [5, 7]. This high community dissimilarity, together with the uneven spatial distribution of α -diversity, also suggests that sufficient sampling efforts (e.g., number and spatial coverage of sampling sites) should be paid to correctly evaluate the microbial (both eukaryotes and bacteria) diversity in cold-seep sediment. We attributed this local divergence of microeukaryotic communities to the high number (or proportion) of endemic species especially at active cold-seep sediment, similar to the findings of prokaryotes at methane seeps [5]. We confirmed this result by showing that community difference between habitats was much higher using UniFrac distance than Bray-Curtis distance, suggesting that the phylogenetic diversification mainly led to the community dissimilarity between habitats, rather than relative abundance of species. One reasonable explanation could be linked to the enhanced parasitism at seep microeukaryotic communities where host communities were largely different (e.g., high metazoan β -diversity, Fig. S7 & S8a) and further selected parasites specifically and influenced other protists differently, ultimately leading to a strong locally shaped biogeographic pattern which occurred in symbiotic bacteria as well [51].

4.2 Mechanisms shaping cold-seep sediment as hotspots for microeukaryotic diversity

This study provides the first empirical evidence thin cold-seep sediment promotes microeukaryotic diversity at a regional level (i.e., γ -diversity), severing as a hotspot for biodiversity. This was reflected by the comparison of total observed ASV richness between habitats from an equal number of randomly extracted sequence or sample in our data. We found that habitat type was more important than depth on shaping microeukaryotic community structure (Fig. S4), which is in accord to the results from a previous study [18] and allows us to treat the whole samples within a habitat as one metacommunity to compare its diversity (e.g., average α - and β -diversity) with other habitats without regard to the depth effects. Our result, showing thin cold-seep sediment had higher microeukaryotic richness than non-seep area, is in accord to the previous studies of microeukaryotes and macrofauna [19, 52], but different from bacterial communities which showed similar richness between cold seep and non-seep regions or moderate level of richness among different habitat types [5, 7]. We attribute this later discrepancy to the different dominant processes structuring the communities of microeukaryotes and bacteria under the same environment, which has been well reported in lakes, seawater and soil [23, 36, 53, 54], but unclear at marine sediment.

In addition to compare microeukaryotic diversity between habitats, our study also explored the main factors controlling them. Habitat quality, including both abiotic and biotic factors, can significantly affect α -diversity [55]. Previous studies showed that variations in some chemical factors can affect the community composition of microeukaryotes at some types of reducing habitats such as oxygen depleted zones and hydrothermal vents [56, 57]. These seep-specific chemical factors, e.g., methane, sulfide and sulfate, can promote the coexistence of diverse bacteria species [58], but their effects on local microeukaryotic richness are unclear. In our study, we found effects from abiotic factors on microeukaryotic α -diversity were not significant. This weak correlation could be explained by the indirect relations between microeukaryotes and the measured chemical factors, for instance, phagotrophic groups (e.g., Amoebozoa, ciliates, MASTs and Retaria) are mainly relied on bacterial food, and parasitic groups (e.g., Apicomplexa and Syndiniales) are determined by their hosts [11, 12]. This was supported by the significant correlations between microeukaryotic richness and the richness of bacteria ($R^2 = 0.22$, $p < 0.001$, by GAM) and metazoan ($R^2 = 0.57$, $p < 0.001$, by GAM) (Fig. 4b, c). Although our study captured some of the most variable chemical factors at seep regions, we acknowledge that other

environmental factors, such as oxygen [59] and organic resources (important to fungi) [60], may largely control the microeukaryotic richness but was unmeasured in our study.

Local (α) diversity can be positive to dispersal rate of species when dispersal is at low to intermediate levels. This is because moderate to intermediate dispersal rates between local communities can weaken local competitive exclusion either by colonization-competition trade-off or source-sink dynamics when resource availability is spatially distributed [34, 61]. Thus, dispersal of species can help maintain high local diversity and stabilize community by increasing species exchange and reducing species distinction caused by local environmental selection, which is called “rescue effects” [62, 63]. This has been tested in marine microbial communities, showing that even a slight increase in dispersal rate can lead to significant changes in microbial richness and community structures [64, 65]. Our study showed that dispersal rate in cold-seep sediment, which was at low to intermediate levels [34], was higher than that at non-seep area from both vertically (between sediment cores) and horizontally (between entire habitats), having positive effects on local microeukaryotic richness. Our results support the idea that cold seeps are island-like habitats harboring diverse endemic species and unique community structures compared with nearby; however, we point out that local microeukaryotic richness here was mainly regulated by dispersal rate of species, while previous studies showed that bacterial richness was locally shaped by some environmental factors (e.g., CH_4) [5, 7]. While in the water column microorganisms’ dispersal is passively favored by ocean currents or disturbances, we argued that the elevated dispersal rate of microeukaryotes in cold-seep sediment could be from the following reasons: (1) parasitic groups, such as Apicomplexa which were more enriched at seep area, can disperse along with metazoan hosts (such as Annelida and Hemichordata with relatively higher sequence abundance in the active seep sediment, Fig. S8) which have strong motility [66]; (2) high sedimentation rate in the Haima cold seep area allows more species exchange of microeukaryotes with sources from water column deposition [67]; (3) seeping gas flow can induce interstitial water circulation, promoting chemical and biological activity (e.g., increasing richness and biomass of nematodes) in cold-seep sediment [20].

The highest β -diversity (i.e., community dissimilarity) contributed to the highest γ diversity of microeukaryotes at ROV1 compared with ROV3 and ROV5, especially when ROV1 and ROV3 had the similar level of α -diversity (richness and Chao1). This finding is in accord to the study of cold seeps in

the Eastern Mediterranean Sea showing that bacterial communities within seep regions varied considerably on spatial scales of only tens to hundreds of meters (similar to the distance between sediment cores within each ROV in our study) [7]. From the fundamental ecological processes in the metacommunity, both dispersal limitation (from neutral theory) and heterogeneous selection (from niche theory) can increase β -diversity [29, 30]. Relative importance of stochastic and deterministic processes could vary greatly and are determined by different environmental conditions and traits of studied organisms [24, 68, 69], leading to the inconsistent conclusions of comparing dispersal limitation and environmental selection in sedimental microeukaryotic communities in previous studies [35, 70, 71]. In our study, effects from dispersal limitation should have minor effects on the β -diversity of microeukaryotic communities at seep area, as no pattern of distance-decay pattern of similarity was observed, which is also supported by the higher dispersal rates calculated at seep area than non-seep area. This is in accord to a previous study showing that microbial communities between seeps can be hardly explained by geographic distances [5]. Further, we found a remarkable increase of heterogeneous selection in the community assembly of microeukaryotes at active seep regions, which could explain the highest β -diversity at ROV1. This reflects that active cold seeps are highly heterogeneous systems with diverse abiotic and biotic resources providing niches for microorganisms [7, 72]. However, compared with bacterial community whose difference at cold seeps were controlled by variations in chemical factors, our results showed that changes in microeukaryotic community compositions (i.e., β -diversity) were mainly regulated by metazoan communities rather than abiotic factors, although the significant correlation to sulfide is in accord to a previous study [18]. Moreover, a much higher correlation between microeukaryotic β -diversity was observed using occurrence data than abundance data (Fig. S7), suggesting that phylogenetic difference of metazoan species, rather than their population size, determined the microeukaryotic community compositions. Together with the result that microeukaryotic richness was mainly determined by metazoan richness (explanation = 58.1%, by GAM, Fig. 4b), our study highlights the importance of parasitism formed by microeukaryotes and metazoans and their biotic interactions on regulating microeukaryotic diversity and community structure at cold seeps.

5 Conclusions

We showed that microeukaryotic communities in Haima cold-seep sediment were spatially divergent even at small scales (e.g., centimeters vertically), and the community difference between habitats became

much larger when considering phylogeny, suggesting the local phylogenesis of microeukaryotes. Our results also suggests that sediment of cold seeps can be hotspots for microeukaryotic diversity. From α -diversity aspect, higher microeukaryotic richness in cold sediment than nearby was mainly promoted by biotic factors, especially metazoan richness, and relatively higher dispersal rate of microeukaryotes. It indicates that both niche and neutral theory-based models should be considered to explore and predict microeukaryotic community biogeography at these island-like cold seeps. It also suggests that sediment or seafloor (besides ocean current) can be important places for species dispersal, genetic connection between cold seeps. From β -diversity aspect, metazoan-mediated selection, rather than from abiotic factors, remarkably increased microeukaryotic community heterogeneity, reflecting the key roles of parasitism in determining microeukaryotic spatial distributions at active cold seeps. We acknowledge that samples from global scales and RNA signatures (e.g., ribosomally active communities) would be important for supporting our findings. Overall, these findings would deepen our understanding of how microbial communities are shaped at extreme habitats and provide indications for conservations of deep-sea biodiversity.

Data Availability

Raw sequencing reads for analyses in this study were deposited in online open database of National Center for Biotechnology Information Search database (NCBI) with projection accession number of PRJNA849592 (submission ID: SUB11547907, to be released in June 2023). R scripts used for analysis can be found from the website: <https://github.com/xzhimenghkust/Haima-cold-seep-microeukaryotes-Rscripts>.

Funding

This study was supported by the Hong Kong Branch of Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou) (SMSEGL20SC01), the Key Special Project for Introduced Talents Team of Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou) (GML2019ZD0409) and the Research Grants Council of Hong Kong (16101917 and 16101318). B.C. was supported by Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou) (SMSEGL20SC02) and a Leverhulme Trust Research Project Grant (PRG-2020-389).

Author Contributions

Hongbin Liu, Zhimeng Xu, Jiawei Chen conceived and designed the study. Jiawei Chen collected the samples and conducted the experiments. Zhimeng Xu performed the analyses and drafted the manuscript. Yingdong Li, Erfan Shekarizz, Wenxue Wu and Bingzhang Chen gave comments and advice on the manuscript.

Conflict of Interest

The authors declare no competing interests.

Supporting information

Figure S1 | Vertical distributions of chemical factors at different habitats. ROV1: active cold seep area; ROV3: less active cold seep area; ROV5: non-seep area, as control. Data was from a long push core (0 – 70 cm) for each ROV while microeukaryotic samples were only taken from 0 cm to 25 cm. Environmental factors of other push cores were in Table S1. ppm: parts-per-million. $\delta^{13}\text{C}_{\text{DIC}}$: $\delta^{13}\text{C}$ carbon-dissolved inorganic carbon.

Figure S2 | Community composition of microeukaryotes at different ROVs. Taxon were divided at a group (class or phylum) level.

Figure S3 | Average community composition of protists from different ROVs. (a) Average values of relative proportion of sequence abundance and ASV richness of samples within each ROV; (b) Number of shared and unique ASVs of ROVs.

Figure S4 | Effects of depth and habitat type on protistan community similarity. (a) Non-metric multidimensional scaling (NMDS) of all samples based on community similarity (Bray-Curtis). Each plot represents one sample, with shapes indicating depths and colors indicating habitats (i.e., ROVs); (b) A detail analysis of similarity (ANOSIM) testing the significant effects of factors (depth and habitat) on shaping protistan communities. Test was performed for combined samples marked with dark filled circles below each bar. A larger ANOSIM-R value suggested greater difference among groups of samples (divided by depth or habitat), with significant codes of *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

Figure S5 | Community beta diversity based on Jaccard dissimilarity at each habitat. Significant difference between ROVs was tested by Wilcoxon rank sum test with significance codes: **: $p < 0.01$, ***: $p < 0.001$, n.s.: not significant.

Figure S6 | Factors determined the protistan communities at cold seep. Here both abiotic and biotic factors were analyzed. Linear regression model was used to show the correlations (Multiple R-squared value and F-statistic p-value).

Figure S7 | Metazoan community composition and difference between habitats. Community compositions of metazoan were shown at the class level by proportion of sequence in each sample (left). Their average proportions in each ROV (i.e., mean relative abundance) were shown in the barchat (right) (a). Difference of metazoan community composition (at the ASV level) among ROVs were compared, with sequence abundance weighted and unweighted, respectively. Community similarity was calculated

High microeukaryotic diversity in the cold-seep sediment

by Bray-Curtis distance and visualized by nMDS with first two dimensions. Each dot represents a sample, with size representing richness and color indicating habitat (ROV). Significant difference ($p < 0.05$) was tested by ANOSIM with a larger R value indicating higher community difference between ROVs (**b**).

Table S1. Distributions of environmental factors

Table S2. Sample metadata and sequencing information

Table S3. Indicator species of ROV1

Table S4. Indicator species of ROV3

Table S5. Indicator species of ROV5

Table S6. Indicator species of ROV1&3 (compared with ROV5)

Table S7 Fitness to neutral community model (NCM) with comparisons to other models

References

1. Boetius A, Wenzhöfer F (2013) Seafloor oxygen consumption fuelled by methane from cold seeps. *Nat Geosci* 6:725–734. <https://doi.org/10.1038/ngeo1926>
2. Valentine DL, Blanton DC, Reeburgh WS, Kastner M (2001) Water column methane oxidation adjacent to an area of active hydrate dissociation, Eel River Basin. *Geochim Cosmochim Acta* 65:2633–2640
3. Sommer S, Pfannkuche O, Linke P, et al (2006) Efficiency of the benthic filter: Biological control of the emission of dissolved methane from sediments containing shallow gas hydrates at Hydrate Ridge. *Global Biogeochem Cycles* 20:. <https://doi.org/10.1029/2004GB002389>
4. Li Z, Pan D, Wei G, et al (2021) Deep sea sediments associated with cold seeps are a subsurface reservoir of viral diversity. *ISME Journal* 15:2366–2378. <https://doi.org/10.1038/s41396-021-00932-y>
5. Ruff SE, Biddle JF, Tesked AP, et al (2015) Global dispersion and local diversification of the methane seep microbiome. *Proc Natl Acad Sci U S A* 112:4015–4020. <https://doi.org/10.1073/pnas.1421865112>
6. van Dover CL, German CR, Speer KG, et al (2002) Evolution and Biogeography of Deep-Sea Vent and Seep Invertebrates. *Science* (1979) 295:1253–1257
7. Ristova PP, Wenzhöfer F, Ramette A, et al (2015) Spatial scales of bacterial community diversity at cold seeps (Eastern Mediterranean Sea). *ISME Journal* 9:1306–1318. <https://doi.org/10.1038/ismej.2014.217>
8. Vanreusel A, Andersen AC, Boetius A, et al (2009) Biodiversity of cold seep ecosystems along the European margins. *Oceanography* 22:110–127. <https://doi.org/10.2307/24860929>
9. Xu T, Wang Y, Sun J, et al (2021) Hidden historical habitat-linked population divergence and contemporary gene flow of a deep-sea patellogastropod limpet. *Mol Biol Evol* 38:5640–5654. <https://doi.org/10.1093/molbev/msab278>
10. Bar-On YM, Phillips R, Milo R (2018) The biomass distribution on Earth. *Proc Natl Acad Sci U S A* 115:6506–6511. <https://doi.org/10.1073/pnas.1711842115>
11. Burki F, Sandin MM, Jamy M (2021) Diversity and ecology of protists revealed by metabarcoding. *Current Biology* 31:R1267–R1280. <https://doi.org/10.1016/j.cub.2021.07.066>
12. Rodríguez-Martínez R, Leonard G, Milner DS, et al (2020) Controlled sampling of ribosomally

- active protistan diversity in sediment-surface layers identifies putative players in the marine carbon sink. *ISME Journal* 14:984–998. <https://doi.org/10.1038/s41396-019-0581-y>
13. de Vargas C, Audic S, Henry N, et al (2015) Eukaryotic plankton diversity in the sunlit ocean. *Science* (1979) 348:1261605
 14. Takishita K, Kakizoe N, Yoshida T, Maruyama T (2010) Molecular evidence that phylogenetically diverged ciliates are active in microbial mats of deep-sea cold-seep sediment. *Journal of Eukaryotic Microbiology* 57:76–86. <https://doi.org/10.1111/j.1550-7408.2009.00457.x>
 15. Takishita K, Tsuchiya M, Reimer JD, Maruyama T (2006) Molecular evidence demonstrating the basidiomycetous fungus *Cryptococcus curvatus* is the dominant microbial eukaryote in sediment at the Kuroshima Knoll methane seep. *Extremophiles* 10:165–169. <https://doi.org/10.1007/s00792-005-0495-7>
 16. Takishita K, Yubuki N, Kakizoe N, et al (2007) Diversity of microbial eukaryotes in sediment at a deep-sea methane cold seep: Surveys of ribosomal DNA libraries from raw sediment samples and two enrichment cultures. *Extremophiles* 11:563–576. <https://doi.org/10.1007/s00792-007-0068-z>
 17. Nagahama T, Takahashi E, Nagano Y, et al (2011) Molecular evidence that deep-branching fungi are major fungal components in deep-sea methane cold-seep sediments. *Environ Microbiol* 13:2359–2370. <https://doi.org/10.1111/j.1462-2920.2011.02507.x>
 18. Pasulka AL, Levin LA, Steele JA, et al (2016) Microbial eukaryotic distributions and diversity patterns in a deep-sea methane seep ecosystem. *Environ Microbiol* 18:3022–3043. <https://doi.org/10.1111/1462-2920.13185>
 19. Levin LA, Mendoza GF, Grupe BM, et al (2015) Biodiversity on the rocks: Macrofauna inhabiting authigenic carbonate at Costa Rica methane seeps. *PLoS One* 10:. <https://doi.org/10.1371/journal.pone.0131080>
 20. O'Hara SCM, Dando PR, Schuster U, et al (1995) Gas seep induced interstitial water circulation: observations and environmental implications. *Cont Shelf Res* 15:931–948
 21. Cao H, Zhang W, Wang Y, Qian PY (2015) Microbial community changes along the active seepage site of one cold seep in the Red Sea. *Front Microbiol* 6:739. <https://doi.org/10.3389/fmicb.2015.00739>

22. Logares R, Deutschmann IM, Junger PC, et al (2020) Disentangling the mechanisms shaping the surface ocean microbiota. *Microbiome* 8:1–17. <https://doi.org/10.1186/s40168-020-00827-8>
23. Wu W, Lu HP, Sastri A, et al (2018) Contrasting the relative importance of species sorting and dispersal limitation in shaping marine bacterial versus protist communities. *ISME Journal* 12:485–494. <https://doi.org/10.1038/ismej.2017.183>
24. Xu Z, Cheung S, Endo H, et al (2022) Disentangling the ecological processes shaping the latitudinal pattern of phytoplankton communities in the Pacific Ocean. *mSystems* 7:e01203-21
25. Feng D, Qiu J-W, Hu Y, et al (2018) Cold seep systems in the South China Sea: An overview. *J Asian Earth Sci* 168:3–16
26. Dini-Andreote F, Stegen JC, van Elsas JD, Salles JF (2015) Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. *Proc Natl Acad Sci U S A* 112:E1326–E1332. <https://doi.org/10.1073/pnas.1414261112>
27. Nemergut DR, Schmidt SK, Fukami T, et al (2013) Patterns and processes of microbial community assembly. *Microbiology and Molecular Biology Reviews* 77:342–356
28. Stegen JC, Lin X, Konopka AE, Fredrickson JK (2012) Stochastic and deterministic assembly processes in subsurface microbial communities. *ISME J* 6:1653–1664
29. Zhou J, Ning D (2017) Stochastic community assembly: does it matter in microbial ecology? *Microbiology and Molecular Biology Reviews* 81:e00002-17. <https://doi.org/10.1128/MMBR>
30. Vellend M (2010) Conceptual synthesis in community ecology. *Quarterly Review of Biology* 85:183–206. <https://doi.org/10.1086/652373>
31. Chase JM, Myers JA (2011) Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366:2351–2363. <https://doi.org/10.1098/rstb.2011.0063>
32. Hubbell SP (2011) *The unified neutral theory of biodiversity and biogeography (MPB-32)*. Princeton University Press
33. Connor EF, McCoy ED (1979) The statistics and biology of the species-area relationship. *Am Nat* 113:
34. Mouquet N, Loreau M (2003) Community patterns in source-sink metacommunities. *Am Nat* 162:544–557
35. Wu W, Liu H (2018) Disentangling protist communities identified from DNA and RNA surveys

- in the Pearl River–South China Sea Continuum during the wet and dry seasons. *Mol Ecol* 27:4627–4640
36. Logares R, Tesson SVM, Canbäck B, et al (2018) Contrasting prevalence of selection and drift in the community structuring of bacteria and microbial eukaryotes. *Environ Microbiol* 20:2231–2240. <https://doi.org/10.1111/1462-2920.14265>
37. Evans S, Martiny JBH, Allison SD (2017) Effects of dispersal and selection on stochastic assembly in microbial communities. *ISME Journal* 11:176–185. <https://doi.org/10.1038/ismej.2016.96>
38. Stegen JC, Lin X, Fredrickson JK, et al (2013) Quantifying community assembly processes and identifying features that impose them. *ISME J* 7:2069–2079
39. Liang Q, Hu Y, Feng D, et al (2017) Authigenic carbonates from newly discovered active cold seeps on the northwestern slope of the South China Sea: Constraints on fluid sources, formation environments, and seepage dynamics. *Deep Sea Res 1 Oceanogr Res Pap* 124:31–41. <https://doi.org/10.1016/j.dsr.2017.04.015>
40. Li S peng, Wang P, Chen Y, et al (2020) Island biogeography of soil bacteria and fungi: similar patterns, but different mechanisms. *ISME Journal* 14:1886–1896. <https://doi.org/10.1038/s41396-020-0657-8>
41. Perkins AK, Rose AL, Grossart H-P, et al (2021) Oxic and anoxic organic polymer degradation potential of endophytic Fungi from the marine macroalga, *Ecklonia radiata*. *Front Microbiol* 12:
42. Mahé F, de Vargas C, Bass D, et al (2017) Parasites dominate hyperdiverse soil protist communities in Neotropical rainforests. *Nat Ecol Evol* 1:1–8
43. Moreira D, López-García P (2003) Are hydrothermal vents oases for parasitic protists? *Trends Parasitol* 19:556–558
44. Zhang Y, Huang N, Wang M, et al (2021) Microbial Eukaryotes Associated With Sediments in Deep-Sea Cold Seeps. *Front Microbiol* 12:782004. <https://doi.org/10.3389/fmicb.2021.782004>
45. del Campo J, Heger TJ, Rodríguez-Martínez R, et al (2019) Assessing the diversity and distribution of apicomplexans in host and free-living environments using high-throughput amplicon data and a phylogenetically informed reference framework. *Front Microbiol* 2373
46. Rueckert S, Simdyanov TG, Aleoshin V v., Leander BS (2011) Identification of a divergent environmental DNA sequence clade using the phylogeny of gregarine parasites (apicomplexa)

- from crustacean hosts. PLoS One 6:e18163. <https://doi.org/10.1371/journal.pone.0018163>
47. Wakeman KC, Leander BS (2013) Identity of environmental DNA sequences using descriptions of four novel marine gregarine parasites, *Polyplacarium* n. gen. (Apicomplexa), from capitellid polychaetes. Marine Biodiversity 43:133–147. <https://doi.org/10.1007/s12526-012-0140-5>
48. Faith DP (1992) Conservation evaluation and phylogenetic diversity. Biol Conserv 61:1–10
49. Faith DP, Baker AM (2006) Phylogenetic diversity (PD) and biodiversity conservation: some bioinformatics challenges. Evolutionary bioinformatics 2:117693430600200000
50. Auguet JC, Barberan A, Casamayor EO (2010) Global ecological patterns in uncultured Archaea. ISME Journal 4:182–190. <https://doi.org/10.1038/ismej.2009.109>
51. Dick GJ (2019) The microbiomes of deep-sea hydrothermal vents: distributed globally, shaped locally. Nat Rev Microbiol 17:271–283
52. Wang Y, Zhang WP, Cao HL, et al (2014) Diversity and distribution of eukaryotic microbes in and around a brine pool adjacent to the Thuwal cold seeps in the Red Sea. Front Microbiol 5:1–10. <https://doi.org/10.3389/fmicb.2014.00037>
53. Wu W, Huang B (2019) Protist diversity and community assembly in surface sediments of the South China Sea. Microbiologyopen 8:e891. <https://doi.org/10.1002/mbo3.891>
54. Powell JR, Karunaratne S, Campbell CD, et al (2015) Deterministic processes vary during community assembly for ecologically dissimilar taxa. Nat Commun 6:1–10. <https://doi.org/10.1038/ncomms9444>
55. Schrader J, Moeliono S, Keppel G, Kreft H (2019) Plants on small islands revisited: the effects of spatial scale and habitat quality on the species–area relationship. Ecography 42:1405–1414. <https://doi.org/10.1111/ecog.04512>
56. Orsi W, Edgcomb V, Jeon S, et al (2011) Protistan microbial observatory in the Cariaco Basin, Caribbean. II. Habitat specialization. ISME Journal 5:1357–1373. <https://doi.org/10.1038/ismej.2011.7>
57. Edgcomb VP, Kysela DT, Teske A, et al (2002) Benthic eukaryotic diversity in the Guaymas Basin hydrothermal vent environment. Proceedings of the National Academy of Sciences 99:7658–7662
58. Meier D v., Pjevac P, Bach W, et al (2017) Niche partitioning of diverse sulfur-oxidizing bacteria at hydrothermal vents. ISME Journal 11:1545–1558. <https://doi.org/10.1038/ismej.2017.37>

59. Zhu P, Wang Y, Shi T, et al (2018) Genetic diversity of benthic microbial eukaryotes in response to spatial heterogeneity of sediment geochemistry in a mangrove ecosystem. *Estuaries and coasts* 41:751–764
60. Orsi WD, Richards TA, Francis WR (2017) Predicted microbial secretomes and their target substrates in marine sediment. *Nat Microbiol* 3:32–37. <https://doi.org/10.1038/s41564-017-0047-9>
61. Matthiessen B, Mielke E, Sommer U (2010) Dispersal decreases diversity in heterogeneous metacommunities by enhancing regional competition. *Ecology* 91:2022–2033
62. Warren PH (1996) The effects of between-habitat dispersal rate on protist communities and metacommunities in microcosms at two spatial scales. *Oecologia* 105:132–140
63. Brown JH, Kodric-Brown A (1977) Turnover rates in insular biogeography: effect of immigration on extinction. *Ecology* 58:445–449
64. Zha Y, Berga M, Comte J, Langenheder S (2016) Effects of dispersal and initial diversity on the composition and functional performance of bacterial communities. *PLoS One* 11:e0155239. <https://doi.org/10.1371/journal.pone.0155239>
65. Shen D, Langenheder S, Jürgens K (2018) Dispersal modifies the diversity and composition of active bacterial communities in response to a salinity disturbance. *Front Microbiol* 9:2188. <https://doi.org/10.3389/fmicb.2018.02188>
66. Danovaro R, Snelgrove PVR, Tyler P (2014) Challenging the paradigms of deep-sea ecology. *Trends Ecol Evol* 29:465–475. <https://doi.org/10.1016/j.tree.2014.06.002>
67. Zhao Z, Sun Z, Wang Z, et al (2015) The high resolution sedimentary filling in Qiongdongnan Basin, Northern South China Sea. *Mar Geol* 361:11–24. <https://doi.org/10.1016/j.margeo.2015.01.002>
68. Yeh YC, Peres-Neto PR, Huang SW, et al (2015) Determinism of bacterial metacommunity dynamics in the southern East China Sea varies depending on hydrography. *Ecography* 38:198–212. <https://doi.org/10.1111/ecog.00986>
69. Zhang H, Huang X, Huang L, et al (2018) Microeukaryotic biogeography in the typical subtropical coastal waters with multiple environmental gradients. *Science of the Total Environment* 635:618–628
70. Kong J, Wang Y, Warren A, et al (2019) Diversity distribution and assembly mechanisms of

- planktonic and benthic microeukaryote communities in intertidal zones of southeast Fujian, China. *Front Microbiol* 10: <https://doi.org/10.3389/fmicb.2019.02640>
71. Zhao F, Filker S, Xu K, et al (2020) Microeukaryote communities exhibit phyla-specific distance-decay patterns and an intimate link between seawater and sediment habitats in the Western Pacific Ocean. *Deep Sea Res 1 Oceanogr Res Pap* 160:103279. <https://doi.org/10.1016/j.dsr.2020.103279>
 72. Cordes EE, Cunha MR, Galéron J, et al (2010) The influence of geological, geochemical, and biogenic habitat heterogeneity on seep biodiversity. *Marine Ecology* 31:51–65. <https://doi.org/10.1111/j.1439-0485.2009.00334.x>
 73. Jens Seeberg-Elverfeldt, Michael Schlüter, Tomas Feseker, Martin Kölling (2005) Rhizon sampling of porewaters near the sediment-water interface of. *Limnology and oceanography: Methods*, 2005, 3(8): 361-371 3:361–371
 74. Elwood HJ, Olsen GJ, Sogin ML (1985) The small-subunit ribosomal RNA gene sequences from the hypotrichous ciliates *Oxytricha nova* and *Stylonychia pustulata*. *Mol Biol Evol* 2:399–410
 75. Bolyen E, Rideout JR, Dillon MR, et al (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 37:852–857
 76. Quast C, Pruesse E, Yilmaz P, et al (2012) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41:D590–D596
 77. Team RC (2020) R: A language and environment for statistical computing (R Version 4.0. 2 R Foundation for Statistical Computing, Vienna, Austria). <https://www.R-project.org/>
 78. Oksanen J, Blanchet FG, Kindt R, et al (2013) Package ‘vegan.’ Community ecology package, version 2:1–295
 79. Hamilton NE, Ferry M (2018) Ggtern: Ternary diagrams using ggplot2. *J Stat Softw* 87:1–17. <https://doi.org/10.18637/jss.v087.c03>
 80. Schneider LK, Anestis K, Mansour J, et al (2020) A dataset on trophic modes of aquatic protists. *Biodivers Data J* 8:1–13. <https://doi.org/10.3897/BDJ.8.e56648>
 81. Flynn KJ, Mitra A, Anestis K, et al (2019) Mixotrophic protists and a new paradigm for marine ecology: Where does plankton research go now? *J Plankton Res* 41:375–391. <https://doi.org/10.1093/plankt/fbz026>
 82. de Caceres, Florian Jansen, de Caceres M M (2016) Package “indicspecies.” indicators 8:

83. Chao A (1984) Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics* 11:265–270
84. Hardy OJ, Senterre B (2007) Characterizing the phylogenetic structure of communities by an additive partitioning of phylogenetic diversity. *Journal of Ecology* 95:493–506. <https://doi.org/10.1111/j.1365-2745.2007.01222.x>
85. Legendre P, de Cáceres M (2013) Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecol Lett* 16:951–963
86. Lozupone C, Lladser ME, Knights D, et al (2011) UniFrac: An effective distance metric for microbial community comparison. *ISME Journal* 5:169–172. <https://doi.org/10.1038/ismej.2010.133>
87. Gotelli NJ (2001) Research frontiers in null model analysis. *Global ecology and biogeography* 10:337–343
88. Wood S (2012) mgcv: Mixed GAM Computation Vehicle with GCV/AIC/REML smoothness estimation
89. Sloan WT, Woodcock S, Lunn M, et al (2007) Modeling taxa-abundance distributions in microbial communities using environmental sequence data. *Microb Ecol* 53:443–455. <https://doi.org/10.1007/s00248-006-9141-x>
90. Sloan WT, Lunn M, Woodcock S, et al (2006) Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environ Microbiol* 8:732–740
91. Burns AR, Stephens WZ, Stagaman K, et al (2016) Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. *ISME Journal* 10:655–664. <https://doi.org/10.1038/ismej.2015.142>
92. Kembel SW, Cowan PD, Helmus MR, et al (2010) Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26:1463–1464. <https://doi.org/10.1093/bioinformatics/btq166>