

Environmental conditions drive zooplankton community structure in the deep-water region of the southern Gulf of Mexico: a molecular approach

Francesco Cicala¹, María Arteaga¹, Sharon Herzka¹, Miguel Martinez¹, Clara Hereu¹, Sylvia Jimenez-Rosenberg¹, Anaid Saavedra-Flores¹, Javier Robles-Flores¹, Ricardo Gomez¹, Paola Batta-Lona¹, and Clara Galindo¹

¹CICESE

May 20, 2021

Abstract

Zooplankton play a pivotal role in sustaining the majority of marine ecosystems. The distribution patterns and diversity of zooplankton provide key information for understanding the functioning of these ecosystems. Nevertheless, due to the numerous cryptic and sibling species and the lack of diagnostic characteristics for immature developmental stages, the identification of the global-to-local patterns of zooplankton biodiversity and biogeography remains a challenge in different research fields. Here, the spatial and temporal changes in the zooplankton community from the open waters of the southern section of the Gulf of Mexico were assessed using a multilocus sequence analysis and metabarcoding approach based on the genetic information of 18S and cytochrome oxidase c subunit I (COI) genes. Additionally, a multi-scale analysis was implemented to evaluate which environmental predictors may explain the variability in the structure of the zooplankton community. Our finding suggests that the synergistic effects of oxygen, temperature, and longitude (intended as a proxy for still unexplored forces) may explain both spatial and temporal changes in the zooplankton community. Furthermore, the zooplankton distribution likely reflects the coexistence of three heterogeneous ecoregions and a bio-physical partitioning of the studied area. Finally, some taxa were either exclusive or predominant with either 18S or COI data. This may suggest that comprehensive assessments of the zooplankton community may be more accurately met by the use of multi-locus approaches.

INTRODUCTION

Zooplankton form some of the most abundant and diverse biological communities in aquatic environments and constitute an integral component of both marine and freshwater ecosystems (Bucklin, Lindeque, Rodriguez-Ezpeleta, Albaina, & Lehtiniemi, 2016; Djurhuus et al., 2018; Xiong et al., 2019). Estimating the richness of zooplankton species remains a priority in several fields of research, from ecology to conservation biology (Carugati, Corinaldesi, Dell'Anno, & Danovaro, 2015; Gaston, 2009). Such studies are particularly important in marine ecosystems where ecological processes are often maintained by both complex biotic interactions and environmental drivers (Palumbi et al., 2009; Steinberg & Landry, 2017). For example, zooplankton communities play a central role in regulating biogeochemical cycles by transferring carbon from primary producers to higher trophic levels (Bucklin et al., 2019; Everaert, Deschutter, De Troch, Janssen, & De Schampelaere, 2018; Lindeque, Parry, Harmer, Somerfield, & Atkinson, 2013; Steinberg & Landry, 2017). Moreover, due to their small size, limited capacity for self-dispersal, and high sensitivity to environmental change, zooplankton are also considered to be useful bio-indicators that may be used to evaluate the health of marine ecosystems (Buttay, Miranda, Casas, González-Quirós, & Nogueira, 2015; Johnston, Mayer-Pinto, & Crowe, 2015; Parmar, Rawtani, & Agrawal, 2016; Yang & Zhang, 2019). Long-term studies have reported that both environmental change and ecological processes shape the spatial and temporal abundance of zooplankton

and the taxonomic composition of their communities (Buttay et al., 2015). As such, environmental change may lead to the loss of zooplankton biodiversity (Gazonato Neto, Silva, Saggio, & Rocha, 2014; Parmar et al., 2016), which in turn may affect ecosystem services and result in economic consequences (Beaugrand, Edwards, & Legendre, 2010; Bucklin et al., 2016; Everaert et al., 2018; Johnston et al., 2015).

The Gulf of Mexico (GoM) is an example of a system that is subject to changing environmental conditions. For example, the large scale near-surface circulation in the GoM is largely dominated by the energetic Loop Current (LC; Damien et al., 2018). The northward penetration of the LC is often accompanied by the formation and release of anticyclonic LC eddies (LCEs), which travel towards the western boundary of the GoM (Damien et al., 2018; Sheinbaum, Athié, Candela, Ochoa, & Romero-Arteaga, 2016). Together with other smaller cyclonic and anticyclonic eddies that are formed within the gulf, LCEs are considered to constitute the principal source of mesoscale variability within this ecosystem (Damien et al., 2018; Hamilton, 2007; Jouanno et al., 2016). Moreover, LCEs and smaller eddies have been found to constrain the distributions of nutrients, zooplankton, and other planktonic species in the gulf (Biggs & Ressler, 2001; Linacre et al., 2015). Based on the distribution of chlorophyll (Damien et al., 2018; Muller-Karger et al., 2015; Salmerón-García, Zavala-Hidalgo, Mateos-Jasso, & Romero-Centeno, 2011), the GoM may be conceptually divided into two main areas: (1) a central oligotrophic area and (2) the eutrophic inshore waters that semi-surround it (Damien et al., 2018; Uribe-Martínez, Aguirre-Gómez, Zavala-Hidalgo, Ressler, & Cuevas, 2019). Nevertheless, few studies have evaluated plankton distributions in the eutrophic inshore waters of the GoM (Biggs & Ressler, 2001; Damien et al., 2018) nor the distribution patterns of zooplankton and other planktonic species over the entire gulf. This constitutes the principle limitation in determining if the zooplankton distribution within the GoM also follows the pattern that has been established for chlorophyll.

The main challenges in evaluating the spatial and temporal patterns of zooplankton communities lie in the taxonomic complexity of their assemblages, which include a considerable number of morphologically diverse and cryptic species and a lack of diagnostic characteristics for immature and larval developmental stages (Bucklin et al., 2016). In recent years, molecular methods have generated new and powerful approaches for assessing zooplankton biodiversity by overcoming the main limitations associated with traditional taxonomic surveys, such as the amount of time needed to identify specimens and the need for advanced taxonomic expertise (Creer et al., 2016; Cristescu, 2014; Zhang, Chain, Abbott, & Cristescu, 2018). In particular, the metabarcoding approach (Creer et al., 2016) is considered to be one of the most comprehensive means to holistically evaluate zooplankton assemblages, as it combines DNA barcoding and high-throughput sequencing to evaluate the taxonomic composition of a sample with any source of environmental DNA (eDNA; Stefanni et al., 2018; Zhang et al., 2018). However, metabarcoding results strongly depend on having adequate taxonomic coverage and a systematic resolution of the chosen molecular marker (Bucklin et al., 2016; Larke, Beard, Swadling, & Deagle, 2017; Zhang et al., 2018).

Nuclear-encoded ribosomal RNA fragments, especially hypervariable regions of the 18S rRNA gene, were prime targets in early studies because they are able to provide conserved primer binding sites with broad taxonomic coverage across the eukaryotic domain (Larke et al., 2017; Lindeque et al., 2013). Nevertheless, the 18S rRNA gene often lacks the taxonomic resolution of protein-coding genes, such as mitochondrial cytochrome oxidase c subunit I (COI; Larke et al., 2017; Machida, Leray, Ho, & Knowlton, 2017; Zhang et al., 2018). Indeed, COI undergoes faster rates of evolution than that of the 18S rRNA gene; hence, its great genetic variability can facilitate the systematic classification of even closely related taxa (Zhang et al., 2018). Nevertheless, the wobble effect (i.e., meaningless nucleotide change in the position of the third codon) can increase the occurrence of primer mismatches among taxa and consequentially induce taxonomic bias during PCR amplification (Larke et al., 2017; Piñol, Mir, Gomez-Polo, & Agustí, 2015). These gene-related limitations have led many researchers to propose the need to implement a multi-locus approach in metabarcoding surveys, as the synergistic information of different loci may represent a trade-off between increasing the resolution needed to identify metazoan taxa (Carroll et al., 2019) and reducing the occurrence of false negative and/or false positive outcomes (Bucklin et al., 2016; Zhang et al., 2018). Accordingly, it has become increasingly accepted that the combined use of nuclear and mitochondrial genes may facilitate the assessment of zooplankton communities due to the resulting improved sensitivity for detecting cryptic

as well as intra-species genetic diversity (Carugati et al., 2015).

Despite the ecological importance of zooplankton and an improved understanding of the dynamics of the GoM, few efforts have been made to describe the composition and distribution of the entire zooplankton community that is dispersed throughout the deep-water region of the gulf in the Exclusive Economic Zone (EEZ) of Mexico. Moreover, the extent to which the physical environment may shape the structure of those communities remains uncertain. In this study, we hypothesized that the structure of the zooplankton community would reflect regional and seasonal environmental features. Accordingly, the main goals of this study were to (i) assess the spatial and temporal variability of the zooplankton community from the deep water region of the GoM within the Mexican EEZ, (ii) explore the possibility that structural changes in the zooplankton community might be related to environmental factors, and (iii) describe the regional and temporal diversity of the community.

In order to address these hypotheses, we characterized the structure and variability of the zooplankton community using a metabarcoding approach based on the genetic information of both the hypervariable V9 and Leray_Folmer 1 regions of the 18S rRNA and COI genes, respectively. We chose this combination of genes because it currently provides the best compromise between the detection and resolution of zooplankton taxa. Specifically, the hypervariable V9 region may be used to amplify across all known and unknown metazoan phyla due to its highly conserved flanking regions (Amaral-Zettler, McCliment, Ducklow, & Huse, 2009). In contrast, we posited that COI could provide greater taxonomic resolution (or α -biodiversity) due to its high variability (Heimeier, Lavery, & Sewell, 2010). For both molecular markers, detected amplicon sequence variants (ASVs) were used to estimate zooplankton alpha and beta diversity across the various sampled areas of the GoM. First, we analyzed the genetic information from three cruises, and we pooled all data from each locus to evaluate the potential zooplankton patterns in a comprehensive analysis. Our efforts were directed towards investigating the effects of temperature, salinity, oxygen, depth, latitude, longitude, water density, and florescence (as a proxy for chlorophyll abundance) in shaping the structure of the zooplankton community in order to better understand the physical-biological interactions controlling spatial, seasonal, and inter-annual changes of zooplankton assemblages.

METHODS

Sample collection

Environmental and biological data were collected from 125 stations covering the EEZ of Mexico in the deep water region of the GoM during three oceanographic campaigns led by Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE; Fig. 1). The XIXIMI-04 campaign was held from 27 August to 15 September 2015 ($n = 47$ stations), while the XIXIMI-05 campaign took place from 10–24 June 2016 ($n = 33$ stations), and the XIXIMI-06 campaign was conducted from 15 August to 8 September 2017 ($n = 45$ stations). The sampling stations were organized latitudinally from $\sim 25^\circ\text{N}$ (Line A) to $\sim 20^\circ\text{N}$ (Line J). Abiotic data and zooplankton samples were collected from 3–8 stations per line. The number of stations sampled during each oceanographic campaign, their geographic positions, and geo-physical parameters are reported in Supplementary Information Table S1.

At each station, zooplankton were collected with a 3m-long bongo net (0.60-m diameter mouth, 335- μm mesh) with a double oblique tow spanning the surface to 200 m. The content of one net was split in two parts. From the first portion, 4/5 of the sample was immediately preserved in 96% ethanol and stored at 4 $^\circ\text{C}$ until further analysis. Continuous measurements of temperature ($^\circ\text{C}$), salinity (psu), dissolved oxygen (DO; ml/L), chlorophyll fluorescence (fru), and pressure (Db) were recorded with a SeaBird SBEplus9 CTD (Sea-Bird Electronics, Bellevue, USA) equipped with calibrated sensors. Raw environmental data were processed with SBE Data Processing software (Seasoft V2 Software Suite, 2013; Sea-Bird Electronics). Given that bongo sampling covered the upper portion of the water column, the average temperature, salinity, DO, fluorescence, and density in the upper 200 m were calculated and used in statistical analyses.

DNA extraction and PCR amplification

Once in the laboratory, DNA extraction was carried out using the protocol proposed by Corell and Rodríguez-Ezpeleta (2013) with modified durations and incubation step temperatures. Briefly, specimens were initially rinsed with an abundant amount of sterile water in order to remove ethanol and other inorganic compounds. Approximately 650 mg of the washed sample was transferred to 1.5-ml sterile microcentrifuge tubes and centrifuged at 3,500 x *g* for 10 min to physically separate the biological material from the residual alcohol supernatant. The DNA was extracted using the sodium dodecyl sulfate (SDS) chloroform method, which involved the following sequential steps. (i) The biological tissues were macerated in 1 ml of SDS buffer [Tris-HCl (10 mM), EDTA (100 mM, pH 8.0), NaCl (200 mM) and SDS (1 %)] until sample homogenization. (ii) The macerated samples were digested with proteinase K at 20 mg/ml (Sigma Aldrich, St. Louis, USA) for 4 h at 65 °C. (iii) The digested samples were centrifuged at 7,500 x *g* for 15 min, and the supernatant was decanted into new 2-ml sterile microcentrifuge tubes containing phenol:chloroform:isoamyl alcohol (25:24:1), followed by gentle mixing and centrifugation at 12,000 x *g* for 15 min. This washing step was repeated twice. (iv) The remaining phenol traces were removed by mixing the sample with chloroform:isoamyl alcohol (24:1) and centrifuging at 12,000 x *g* for 10 min before transferring the DNA supernatant to new 2-ml sterile microcentrifuge tubes. (v) DNA precipitation was achieved by adding 95% ethanol and 3 M sodium acetate and incubating the samples at -80 °C for 1 h. (vi) After centrifugation at 12,000 x *g* for 20 min, the remaining salts were removed with 200 µL of 80% ethanol. This step was repeated twice. (vii) Ethanol was removed by decanting, and the DNA pellet was dried at room temperature for ~10 min. (viii) Purified DNA was re-suspended in 100 µl of Milli-Q water and stored at -20 °C until further analysis.

DNA quantity and purity were evaluated with a Qubit[®] 3.0 fluorimeter using a Qubit[®] dsDNA Broad Range assay (Life Technologies, Grand Island, USA) and a Nanodrop 2000 spectrophotometer (Thermo Scientific, Waltham, USA), respectively. Finally, DNA integrity was assessed by electrophoresis and visualized in a 1% agarose gel. To sequence the extracted DNA with Illumina MiSeq technology, we implemented a dual PCR amplification method. During the first PCR round, we amplified the hypervariable V9 region (~130 bp) of small ribosomal 18S and the Leray-Folmer 1 region (~313 bp) of the COI gene. These fragments were amplified with the use of the universal eukaryote primers 1389F (5'-TTGTACACACCGCCC-3') and 1510R (5'- CCTTCYGCAGGTTACACTAC-3'; Amaral-Zettler et al., 2009) and mlCOIintF (5'-GGWACWGGWTGAACWGTWTAYCCYCC-3') and jgHCO2198 (5'-TAIACYTCIGGRTGICCRARAAYCA-3'; Leray et al., 2013) for 18S rRNA and COI, respectively.

18S rRNA was amplified with KAPA HiFi HotStart DNA Polymerase (KAPA Biosystems, Basel, Switzerland). Each reaction mix contained 1x KAPA HiFi Buffer GC, 0.3 mM of KAPA dNTP Mix, 0.20 µM of each primer, 1 U of HiFi HotStart DNA Polymerase (KAPA Biosystems), and 150 ng of purified DNA in a total reaction volume of 23 µL. Thermal cycling consisted of an initial incubation at 95 °C for 3 min, followed by 25 cycles of 10 s at 98 °C, 30 s at 68 °C, and 30 s at 72 °C, with a final extension of 10 min at 72 °C. The PCR products were visualized in a 2% agarose gel and then purified with Agencourt AMPure XP paramagnetic beads (Beckman Coulter, San Diego, USA).

The COI marker was amplified with a Platinum Taq DNA Polymerase (High Fidelity; Invitrogen, Carlsbad, USA) in a reaction mix containing 1x High Fidelity Buffer, 3.0 mM of MgSO₄, 0.2 mM of dNTP Mix, 0.20 µM of each primer, 1 U of Platinum Taq, and 20 ng of purified DNA in a total reaction volume of 20 µL. Thermal cycling consisted of an initial incubation at 94 °C for 3 min, followed by 38 cycles of 30 s at 94 °C, 30 s at 46 °C, and 90 s at 72 °C, with a final extension of 5 min at 72 °C. The PCR products were visualized in a 2% agarose gel and then purified with Agencourt AMPure XP paramagnetic beads (Beckman Coulter).

In the second PCR, we incorporated the Illumina index using the Nextera XT kit following the instructions of the manufacturer using a high fidelity polymerase (KAPA HiFi HotStart ready mix, KAPA Biosystems). The indexing amplification program included 3 min at 95 °C, followed by 12 and 8 cycles (for the 18S and COI primers, respectively) of 30 s at 95 °C, 30 s at 60 °C, and 30 s at 72 °C, and a final extension of 5 min at 72 °C. The PCR products were visualized in a 2% agarose gel before being cleaned with two rounds of AMPure XP magnetic beads (Beckman Coulter). The first and second purification rounds targeted nonspecific product sizes and the sizes of expected amplicons, respectively. The cleaned PCR results were

visualized in 2% agarose gels and quantified with a Qubit? 3.0 fluorometer using a Qubit? dsDNA High Sensitivity assay (Life Technologies) before normalization to 40 nM. Finally, each library was consecutively diluted until reaching 4 pM and sequenced with 30% PhiX (Illumina) in an Illumina MiSeq machine at 2 x 150 bp and 2 x 300 bp configurations for the 18S and COI markers, respectively.

Amplicon library preparation and bioinformatic analyses

Illumina outputs were demultiplexed using Illumina MiSeq control software (v.2.6.2.1). The raw DNA libraries were deposited at the National Center for Biotechnology Information (NCBI) under the BioProject ID PRJNA630297. The primers were removed using cutadapt v.1.15 (Martin, 2011). Additionally, traces of both primers and Illumina NexteraPE-PE adaptors were removed using trimmomatic v. 0.38 (Bolger, Lohse, & Usadel, 2014). Bases with average quality scores below 20 with a sliding window of 4 bases were trimmed. Trimmed reads were then imported and analyzed using R (R Core Team, 2014) and Quantitative Insights Into Microbial Ecology v. 2020.6 (QIIME2; Bolyen et al., 2019). Given that the DNA libraries were sequenced during different runs, reads for both loci were initially filtered, denoised, and merged with DADA2 (Callahan et al., 2016) using DADA2 workflow for big data (hosted at https://benjjneb.github.io/dada2/bigdata_1-2.html). The same program was also used to detect and remove chimeras. Sequences were then pooled in ASVs (100% identity), and singletons were discarded. Feature tables were rarefied at the minimum sample read depth using the QIIME2 *diversity alpha-rarefaction* plugin and implemented in a downstream analysis. Amplicon sequence variants were then aligned with MAFFT v.7 (Katoh & Standley, 2013), and a phylogenetic tree was built and rooted with fastTree v.2.1 (Price, Dehal, & Arkin, 2010). Taxonomic assignments were conducted using the naive Bayesian approach *rdp* classifier (Wang, Garrity, Tiedje, & Cole, 2007). The classifier was trained on the BOLD (Ratnasingham & Hebert, 2007) and SILVA v.132 (Quast et al., 2013) databases for COI and 18S rRNA, respectively.

We limited our analyses to known zooplankton taxa and discarded reads that were not identified to the phylum level with > 80% confidence. Accordingly, ASVs with no taxonomic match (NA) or those not assigned to the kingdom Animalia were removed. Moreover, taxonomy lineages from the World Register of Marine Species (WoRMS Editorial Board, 2017) were retrieved with a synonymous per taxa method using the taxize package (Chamberlain & Szocs, 2013) in order to include only marine zooplankton species and to match the obtained organism nomenclature with the SILVA and BOLD databases. Finally, in order to minimize the chance of spurious ASVs being included in the final datasets, the number of reads for each ASV were log transformed and ranked according to read abundance (Supplementary Information Fig. S1). Notably, we observed that ASVs composed of < 12 and < 21 reads for 18S and COI, respectively, were generally detected in only one sample. Thus, we considered those ASVs as PCR amplification and/or sequencing noise, and we removed them from the final datasets.

Ecological analyses

In order to evaluate how exhaustively the zooplankton communities of each station were sampled, rarefaction curves of the detected ASVs were generated using the QIIME2 *diversity alpha-rarefaction* plugin, implementing the median frequency of the reads as the sequencing depth. Additionally, the number of obtained ASVs from each station was compared against the non-parametric species richness estimator Chao1 (Chao, 1984; Supplementary Information Fig. S2). The rarefied number of reads was used as an abundance proxy to estimate the Shannon diversity index and structural changes in the zooplankton community among spatial and temporal scales.

Differences in taxonomic structure among cruises were evaluated at the order level with bar-plots based on only the most abundant ASVs (cutoff at > 1% frequency) and at the ASV level by principal coordinate analysis (PCoA) based on Bray Curtis, Jaccard, and weighted and unweighted phylogenetic Unifrac distance metrics. The results of the PCoA were visualized with EMPERor (Vazquez-Baeza, Pirrung, Gonzalez, & Knight, 2013). Statistical differences were evaluated by a permutational multivariate analysis of variance (PERMANOVA) and were based on 4999 permutations using *beta-group-significance* plugin of QIIME2 v.2019.7 (Anderson, 2001). Also, shared and/or exclusive non-rarefied ASVs among the same stations from

each cruise ($n = 19$ and 21 for 18S and COI, respectively) were visualized using Venn diagrams created with the open web-software InteractiVenn (<http://www.interactivenn.net/>; Heberle, Meirelles, da Silva, Telles, & Minghim, 2015). Significant differences in the zooplankton diversity index were evaluated with a one-way analysis of variance (ANOVA) conducted in Statistica (StatSoft Inc., Tulsa, USA) with statistical significance evaluated at $\alpha = 0.05$ ($p < 0.05$). Finally, to determine which ASVs were primarily contributing to the observed spatial-temporal dissimilarities, a SIMPER analysis based on rarefied ASV abundance was implemented in PRIMER+P v.6 (K. R. Clarke & Warwick, 2001).

Environmental data, zooplankton abundance, and community structure

With the aim of evaluating the correlations between environmental parameters and the structure of the zooplankton community, revealing the spatial structure within each cruise, and revealing seasonal and inter-annual environmental changes, we analyzed the relationships between the 8 environmental predictors and the zooplankton distribution. Those predictors included the average (0-200 m) environmental conditions (i.e., temperature, oxygen, salinity, fluorescence, and density) as well as geographic and bathymetric predictors (e.g., latitude, longitude, and depth). Since the cruises reflect late spring (XIXIMI-05) and summer (XIXIMI-04 and XIXIMI-06) conditions, season was used as a constraining factor in subsequent analyses.

The effect of environmental parameters was tested using the proxy of abundance for each taxa at the family level. A distance-based linear modeling (DistLM) analysis was applied with a multivariate multiple regression analysis using Primer 6+P (K. R. Clarke & Warwick, 2001). We conducted these analysis to estimate the independent ordination of all predictors (Marginal test), which determines the proportion of zooplankton variance explained by each environmental variable independently, and to obtain an optimal ordination model (Sequential Test), in which the model partitions the variation in the data based on a multiple regression model selected by the user (e.g., forward, stepwise, or best fit; K. R. Clarke & Warwick, 2001). The latter outcome may be considered to be the best statistical combination of all abiotic predictors. In this study, we implemented the Akaike Information Corrected Criterion (AICc) and ‘stepwise’ options for model selection. We selected this approach because stepwise multiple regression adds or subtracts predictor variables from a model until most of the variation is explained; variables are excluded if they behave like random variables in terms of the additional variation explained. A significance evaluation of the multidimensional model was conducted with a permutation method with 9999 permutations (Usov, Khaitov, Smirnov, & Sukhotin, 2019) implemented in Primer 6+P (K. R. Clarke & Warwick, 2001).

A distance-based redundancy analysis (dbRDA) was performed using Primer 6+P (K. R. Clarke & Warwick, 2001) to visualize the relative importance of all predictor variables (McArdle & Anderson, 2001). The potential relationships were evaluated using normalized environmental data and the fourth root transformed read abundance of zooplankton at the family level. The fourth root transformation was deemed to be the most appropriate transformation since it reduces the weight of highly abundant taxa and facilitates comparisons among different datasets (Howson, Buchanan, & Nickels, 2017; Vause et al., 2019). The resultant data were converted to a resemblance matrix using the Bray-Curtis similarity index.

According to the multivariate multiple regression results (reported below), stations were categorized into discrete categories according to their environmental profiles. These categories were: low/high oxygen concentration, warm/cold temperatures, and east/west longitude. With regard to the latter, east stations were located at longitudes < 86.30 °W, while the remaining stations (> 86.30 °W) were attributed to the west group. Likewise, stations with lower/higher values than the average mean values for oxygen and temperature (specific average threshold values of oxygen and temperature were calculated according to the stations in each analysis) were grouped as “Low/High- O_2 ” and “Cold/Warm”, respectively. These oceanographic variables (averaged for the first 200 m of the water column) were visualized in contour plots generated with Ocean Data View software (Schlitzer, 2018). This approach was used to compare spatial (within each cruise) and temporal patterns (among spring and summer cruises) and to test for similarities/differences in zooplankton community structure. In addition, all data for the three cruises were also included in a comprehensive analysis.

Potential spatial and seasonal segregation of the environmental parameters after normalizing the data was tested using a Cluster/SIMPROF analysis in Primer 6+P software. Clustering was carried out with a Euclidean distance matrix using the group average method. Statistical differences among groups of stations were evaluated by a PERMANOVA with 4999 permutations using Primer 6+P software, while comparisons of the average abiotic parameters were conducted with ANOVAs in Statistica software.

RESULTS

Sequencing results and overall taxonomic composition of the zooplankton community

A total of 242 (122 and 120 for 18S rRNA and COI, respectively) zooplankton community libraries were successfully amplified. Overall, sequencing produced more than 28.5 million paired-end reads, and 24.4 million sequences (~86% of the total number of reads) met quality control tests. After library curation, an inflection point in ASV abundance was observed at 1.2 and 1.3 log value for 18S rRNA and COI, respectively (Supplementary Information Fig. S1A and S1B). Moreover, ASVs below these threshold values were generally detected for only one station and were considered as *in vitro* local artifacts that likely rose during amplification and/or sequencing. Therefore, further analyses were carried out with ASVs with a minimum number of reads, namely 12 (log value ~1.1) and 21 (log value ~1.3) for 18S rRNA and COI, respectively. The final datasets included 13,088,619 reads representing 1,247 ASVs for 18S and 8,573,204 reads representing 14,795 ASVs for COI (Supplementary Information Table S2 and Table S3). The number of reads and ASVs removed during the quality control tests for both loci are reported in Supplementary Information Table S4.

According to the asymptotic shape of the rarefaction curves (Supplementary Information Fig. S3 and Fig. S4) and the closeness between the detected ASVs to Chao1 (Supplementary Information Fig. S2), zooplankton communities were exhaustively sampled for both genetic markers. The minimum number of reads was 20,076 and 57,585 for the COI and 18S datasets, respectively. Thus, the samples were rarefied to these read numbers.

Taxonomic classification varied somewhat depending on the molecular marker. For example, some taxa were exclusively detected with either 18S rRNA (e.g., Brachyopoda, Bryozoa, Ctenophora, Hemichordata, and Nemertea) or COI (e.g., Chordata and Rotifera). Moreover, we observed a discrepancy in the relative abundance of taxa depending on the locus (e.g., Calanoida and Euphausiacea; Fig. 2A and 2B). Specifically, the most common orders and their relative abundance based on COI and 18S rRNA, respectively, were (abundance percentages in parenthesis): Calanoida (42.5%, 51.1%), Euphausiacea (17.5%, 19.5%), Aphragmophora (9.3%, 2.3%), Halocyprida (8.7%, 1%), Siphonophorae (5.9%, 5.1%), Decapoda (2.6%, 2.1%), and Phyllococida (0.06%, 1.2%) Oegopsida (2%, 0.1%). Additionally, the predominant orders detected only with COI were Myctophiformes (2%), Stomiiformes (1.6%), and Clupeiformes (1.6%), while those detected only with 18S rRNA included Salpida (1.3%) and Doliolida (1.3%).

Relationships between environmental conditions and the zooplankton community

In order to evaluate the extent to which physical environmental variables may shape the spatial and temporal structure of the zooplankton community in the GoM, the relationships between environmental data and taxa-abundance matrices were evaluated through DistLM. These analyses were conducted using both the genetic information of each cruise and in a comprehensive analysis in which all stations (from all cruises) were included. For the latter, the results of the DistLM marginal tests indicated that each of the environmental predictors (when considered in isolation) explained a fraction of the variation observed in the zooplankton community structure (ranging from 2% up to 8.5%; Table 1A and 1B). However, the sequential tests showed that up to 18% of zooplankton diversity may be explained by the synergistic combination of DO, temperature, and longitude for both loci (Table 1C and 1D; $R^2 < 0.19$, $P < 0.0004$). Moreover, for only 18S, an additional 1.68% of zooplankton variability may be explained by salinity (Table 1C). Similar results were also obtained with the individual cruise datasets (Supplementary Information Table S5).

The relationship between environmental parameters and abundance at the family level detected with both loci was visualized with a dbRDA plot. The first axis captured up to 43.1% of the fitted variability and 8.4% of the total variation among both loci and was mainly correlated with two hydrographic variables (oxygen

and temperature; Fig. 3A and 3B). The second axis captured up to 35.2% of the fitted variability and ~6% of the total variation, and the strongest correlation was with longitude. Notably, zooplankton abundance scaled negatively with DO, temperature, and longitude for the majority of stations, indicating that a higher degree of complexity in the zooplankton community could be present in waters that are relatively cold and that have low dissolved oxygen concentrations, which were located in the western portion of the GoM.

In order to further evaluate this result and explore the potential differences in community structure between stations, we clustered each station considering its hydrographic features and geographic position (Fig. 4). Our finding suggests the presence of three different “ecoregions” in the deep water region of the southern GoM. The first ecoregion mostly included stations off the eastern coast of the Yucatan platform (sampling line Y, hereafter Y) located at $< 86.30^\circ\text{W}$ (Fig. 5 and Fig. 6), which were mainly characterized by high oxygen concentrations and temperatures. Moreover, among the western stations of the GoM ($> 86.30^\circ\text{W}$), an additional latitudinal boundary was observed in the proximity of the 22°N parallel (Fig. 6). Indeed, the northern stations (sampling lines: A, B, and C; hereafter N) generally presented higher concentrations of oxygen and higher water temperatures than those of the southern stations (sampling lines: D, E, F, G, H, and J; hereafter S). However, the proposed partition was less clear for the stations located along Lines C, D, and E, which suggests that environmental differences are most notable between the extreme southern and northern sectors of the surveyed area and less marked in between (Fig. 6).

Overall, the hydrographic measurements followed trends that are expected with seasonal change; salinity and temperature were higher during summer (and highest in August-September 2017 cruise XIXIMI-06) while oxygen, fluorescence, and density were higher in late spring (June 2016 cruise XIXIMI-05). However, no significant inter-annual differences were observed with regard to hydrographic predictors with the exception of dissolved oxygen, which was significantly higher in late spring (XIXIMI-05) compared to that in summer (XIXIMI-04 and XIXIMI-06; one-way ANOVA; $F = 9.97$, $p < 0.0002$). Additionally, we observed significant differences in the environmental variables among the N, S, and Y ecoregions (PERMANOVA; pseudo-f > 4.7623 ; $p = 0.0002$).

Spatial and temporal variation of the zooplankton community

The spatio-temporal structure of the zooplankton community was investigated at both the family (mean abundances; SIMPER) and ASV-species (Bray Curtis, Jaccard, weighted and unweighted Unifrac distance metrics; PCoA) taxonomic levels. Over the spatial scales examined, the results of the PCoAs obtained using the genetic information from all stations supported a significant differentiation of the zooplankton community between regions with low/high oxygen concentrations, warm/cold temperatures, and east/west longitudes (Fig. 7 from A to F; PERMANOVA; pseudo-f: > 5.088 ; $p < 0.0004$). Notably, the regional differences in zooplankton community structure were consistent with ecoregion partitioning based on the aforementioned environmental and geographic variables (Fig. 3A and 3B; PERMANOVA; pseudo-f: > 2.346 ; $p < 0.0002$). In addition, higher zooplankton alpha-diversity was found at stations with low mean oxygen concentrations and low water temperatures (Southern ecoregion; Fig. 8A and 8B). Nevertheless, this increase was only statistically supported for zooplankton communities detected with the use of 18S rRNA (average Shannon diversity: N = 4.69, S = 4.93, and Y = 4.15; ANOVA; $F = 13.54$, $p = 0.0001$). Similar results were also observed when the datasets from each expedition were individually analyzed (Supplementary Information Fig. S5 and Fig. S6). Once again, zooplankton community structure varied regionally and was mainly associated with local changes in oxygen and/or local temperature. Finally, the empirical north-south spatial boundary around the 22°N parallel was statistically supported for XIXIMI-04 and XIXIMI-06, whereas an alternative oxygen and/or temperature water column configuration was observed for XIXIMI-05 (Fig. 6). Specifically, during this early summer expedition, warm waters formed at a south-west boundary, resulting in a south-west/north-east regional partitioning of the GoM (Fig. 6).

Given that significant differences in community structure were observed, we implemented a SIMPER analysis in order to determine the contribution of each taxa to the dissimilarity among the groups of samples. The same analysis was conducted for both genetic markers; nevertheless, the discrepancy in taxa abundance among the markers hampered an exhaustive comparison between zooplankton communities detected with

18S rRNA and COI, which was outside the scope of this study. As a result, we cannot propose an unambiguous pattern for all the taxa, but we observed that Centropagidae, Salpidae, Euplotheutidae, Sagittidae, and Pyrosomatidae, among others, were mainly detected at stations with lower than average water temperatures and relatively low oxygen concentrations, which were found in the western GoM. Accordingly, those taxa primarily contributed to the spatial segregation of the southern ecoregion. On the contrary, other taxa, such as Scaridae, Calanidae, and Processidae, mainly contributed to making the Yucatan bio-ecoregion distinct, as they were mostly detected in warm waters with high concentrations of oxygen in the eastern GoM. The taxa contribution (cutoff at 1%) for 18S and COI is reported in Supplementary Information Table S5.

Over the temporal scales analyzed, the dynamics of the zooplankton communities exhibited similar patterns to those mentioned above. In this regard, the PERMANOVA based on all distance metrics revealed significant differences between all three sampling periods (pseudo-f: > 2.177 ; $p < 0.0002$; Fig. 9); nevertheless, those variations were not accompanied by significant changes in zooplankton diversity. Once again, we observed that differences in zooplankton community structure may reflect seasonal environmental conditions, which may consequently promote changes in the abundance of certain taxa, such as Mysidae, Hormathiidae, and Euphausiidae, among others. These taxa were mainly detected in spring (XIXIMI-05) when significantly higher oxygen concentrations were detected. Likewise, during the summer expeditions (XIXIMI-04 and XIXIMI-06), we observed an increase in the abundance of Euphausiidae, Calanidae, and Hormathiidae, which suggests that high water temperatures may be positively associated with these taxa (Supplementary Information Table S6).

DISCUSSION

Partitioning of the GoM based on a multi-locus assessment of the zooplankton community

Both the results of the multivariate regression and the PCA suggest a clear spatial and temporal segregation of zooplankton that was mainly explained by oxygen, temperature, and longitude gradients. The importance of longitude in the model suggests that additional unexplored predictors related to this variable may also play key roles in shaping the composition of the zooplankton community. In this sense, we did not consider the longitudinal spatial gradient as a source of ecological patterns per se but as a proxy of still unrevealed environmental predictors that may explain a portion of the variability in the zooplankton community (Gluchowska et al., 2017; Hawkins & Diniz-filho, 2004). Our results suggest that the physical characteristics of the water column may support the occurrence of at least three heterogeneous ecoregions within the studied area. Such ecoregions comprise the north, south, and an eastern area around the Yucatan peninsula. Furthermore, the characteristics of the water profiles of each region may determine the distribution of at least some zooplankton species, resulting in the aforementioned structural segregation. Accordingly, the main empiric boundaries inferred in this study may be located around 86°W and 22°N , which coincide with the presence of the LC (eastern boundary) and with the southern edge of quasi-permanent cyclonic and anticyclonic eddies (north-south boundary), respectively.

So far, few efforts have described the pelagic communities across the entire GoM; however, our proposed partitioning strongly supports earlier insights based on both chlorophyll concentrations and water mass dynamics. In this context, Damien et al. (2018) proposed the first partitioning of the GoM along 22°N based on chlorophyll concentrations, while Sheinbaum et al. (2016) estimated that among 60–80% of the horizontal variance of the Yucatan Channel is directly and/or indirectly related with the LC. The proposed GoM partitioning was clearly evident during the summer expeditions (cruises XIXIMI-04 and XIXIMI -06), whereas an alternative oxygen-temperature partitioning was observed and related with seasonal water dynamics during XIXIMI-05. The water analysis of this cruise revealed the occurrence of two anticyclonic eddies that crossed 22°N , creating a north-south boundary and willing the GoM into a diagonal south-west/north-east thermal configuration (Fig. 6). As this boundary was detected once, we cannot state if this water organization may be considered a usual season-specific configuration or an anomaly, and more research is needed to resolve this issue.

Finally, the zooplankton in the GoM seems to form a more stable community in the southern region (south of

22 °N) compared to that of the northern region. Indeed, stations from sampling Lines F, G, H, and J showed low structural variability over the 3 years of observations. This is likely explained by the higher productivity observed in this section of the GoM compared to that of the southern region. Indeed, the northern region is under the influence of a quasi-permanent cyclonic gyre, which has been associated with higher nutrient concentrations, fluorescence, and productivity compared to that of the southern region (Färber Lorda, Athié, Camacho Ibar, Daessle, & Molina, 2019; Linacre et al., 2015; Pérez-Brunius, García-Carrillo, Dubranna, Sheinbaum, & Candela, 2013). The northern region is also influenced by upwelled waters from the Bank of Campeche (Salmerón-García et al., 2011).

Temporal zooplankton variability may be also related with oxygen, temperature, and longitude gradients. Although low environmental variability was observed during the study, the abundance of the same taxa, such as Mysidae, Hormathiidae, Euphausiidae, and Calanidae, likely varied according to the co-occurrence of certain environmental conditions. An analysis of the distributions of single taxa was not directly addressed in this paper since our goal was to study the temporal effects of abiotic factors on the entire zooplankton community. However, the integration of single taxa into the analysis is essential to fully understand the functioning of the GoM ecosystem. Their representation in the model is critical and requires a set of taxon-specific DNA libraries that were not available when this study was conducted; however, the development a finer taxonomic scale is currently underway.

Detecting the spatio-temporal variability in the zooplankton community in this study relied on the information of a multi-locus metabarcoding approach. Notably, despite the aforementioned observed taxonomic discrepancies among the two loci employed in this study, similar zooplankton distributions were obtained for both 18S and COI. This suggests that the results obtained strongly reflect the natural variability of the zooplankton community, and methodological bias (e.g., implemented molecular markers) is likely to have only marginally affected the results. It is probable that the main limitation associated with using the combined information of COI or 18S rRNA markers lies in the incompleteness of reference databases and that not all zooplankton species have reference sequences deposited for either the COI or 18S rRNA markers (Larke et al., 2017; Stefanni et al., 2018). However, in contrast to previous studies that recommended 18S rRNA as the most suitable marker for surveying zooplankton communities (Zhan, Bailey, Heath, & Macisaac, 2014), our insights suggests that COI provide similar taxa coverage of zooplankton phyla with higher taxonomic resolution, as suggested by Machida et al. (2017). Finally, our results support that COI may provide better taxa classifications for lower taxonomic levels compared to those of 18S rRNA (Larke et al., 2017; Stefanni et al., 2018). In this study, we observed that 18S rRNA presented a relatively high affinity to Calanoida and Euphausiacea while COI provided relatively uniform taxonomic coverage (L. J. Clarke, Soubrier, Weyrich, & Cooper, 2014; Deagle, Jarman, Coissac, Pompanon, & Taberlet, 2014). Due to the taxonomic complementarity of those loci, we suggest that most comprehensive assessments of zooplankton biodiversity may be conducted using a synergistic multi-locus approach.

Effect of environmental predictors on zooplankton diversity

As previously observed for the northern GoM, environmental conditions in the southern GoM “were comparable among years, but more variable in space” (Elliott, Pierson, & Roman, 2012). Indeed, similar patterns of zooplankton spatial aggregations over-time allowed us to combine the genetic information of each field expedition in order to explore the spatio-temporal variability in the zooplankton community as a whole in a comprehensive analysis.

Our findings suggest that all tested environmental predictors may promote changes in the structure of the zooplankton community, even with low contribution percentages. The simplest explanation for such a poor relationship is that other factors, which were not taken into consideration (e.g., nutrient variation and irradiance, among others), could also be involved in shaping the structure of the zooplankton community of the study area. Moreover, biotic interactions within a zooplankton community, such as predation or symbiotic associations, may also shape the structure of the community (Gasca, Suárez-Morales, & Haddock, 2007; Hereu et al., 2020; Jennifer E. & Mary N., 2001). However, alternative and/or complementary possibilities may be also proposed. In this regard, it has been previously reported that environmental predictors generally show

patchiness (Espinasse, Carlotti, Zhou, & Devenon, 2014; Gluchowska et al., 2017; Trudnowska, Gluchowska, Beszczyńska-Möller, Blachowiak-Samolyk, & Kwasniewski, 2016; Usov et al., 2019), which hampers pattern assessments over a continuous space (Helenius, Leskinen, Lehtonen, & Nurminen, 2017; Usov et al., 2019). This may become particularly relevant in studies that consider large-scale geographic areas or the open ocean, in which environmental gradients usually show three-dimensional reorganizations (Usov et al., 2019). Thus, multiple factors like winds, currents, and water mass movements, among others, may disrupt the linearity of hydrological boundaries (Espinasse et al., 2014; Usov et al., 2019). In our study, the discontinuity of hydrological patterns mainly challenged the cluster organization of the sampling stations. Indeed, even stations that were geographically close to one another presented dissimilarities in environmental profiles that ranged from small to large. These divergences mostly concerned stations located in the central portion of the sampling area (Lines C, D, and E), suggesting that this area may be considered a transitional region separating the northern and southern ecoregions.

Finally, the low strength of the correlations between abiotic and biotic factors may be also explained by considering the experimental design. One of the aims of this study was to relate the environmental variables with the structural patterns of the zooplankton community. Accordingly, the environmental data were evaluated in the context of the zooplankton community, which may have hidden taxa-specific ecological niche partitioning. For example, the abundance of a given taxon may increase in response to a specific combination of environmental conditions, but simultaneously, the same conditions may provoke a decrease in the abundance of other taxa (Elliott et al., 2012). Hence, the observed low environmental contribution percentages may have been the result from opposing taxa-specific responses to the same environmental cues. If validated, this supposition may be used to re-evaluate the implications of the observed correlations. In this context, we posit that the inclusion of additional predictors into the model may not necessarily result in a strengthening of the correlations between biotic and abiotic variables, as some taxa will likely respond positively to a new predictor while others may respond negatively. Nonetheless, additional research is needed to support or reject this hypothesis.

Acknowledgements

This research was funded by the National Council of Science and Technology of Mexico - Mexican Ministry of Energy - Hydrocarbon Trust, project 201441 and Gulf of Mexico Research Consortium (CIGoM). We would like to thank the Ocean Leadership/Gulf of Mexico Research Initiative (GOMRI) [grant number 16-052 to Dr. W. Kelley Thomas; SA 16-18] and the National Institutes of Health New Hampshire Idea Network of Biological Research Excellence [grant number 5 P20 GM 10350605] for supporting this study and bioinformatic training and support. We are grateful to Andrea Lievana-MacTavish for a critical assessment of an earlier version of the manuscript and English editing.

REFERENCES

- Amaral-Zettler, L. A., McCliment, E. A., Ducklow, H. W., & Huse, S. M. (2009). A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA. *Genes. PLoS ONE*, *4* (7), 1–9. doi: 10.1371/journal.pone.0006372
- Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, *26*, 32–46. doi: 10.1080/13645700903062353
- Beaugrand, G., Edwards, M., & Legendre, L. (2010). Marine biodiversity, ecosystem functioning, and carbon cycles. *Proceedings of the National Academy of Sciences of the United States of America*, *107* (22), 10120–10124. doi: 10.1073/pnas.0913855107
- Biggs, D. C., & Ressler, P. H. (2001). Distribution and abundance of phytoplankton, zooplankton, ichthyoplankton, and micronekton in deepwater Gulf of Mexico. *Gulf of Mexico Science*, *19* (1), 7–29. doi: 10.18785/goms.1901.02
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30* (15), 2114–2120. doi: 10.1093/bioinformatics/btu170

- Bolyen, E., Rideout, J., Dillon, M., Bokulich, N., Abnet, C., Al-Ghalith, G., ... Caporaso, G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology* , 37 (August), 852–857.
- Bucklin, A., Lindeque, P. K., Rodriguez-Ezpeleta, N., Albaina, A., & Lehtiniemi, M. (2016). Metabarcoding of marine zooplankton: Prospects, progress and pitfalls. *Journal of Plankton Research* , 38 (3), 393–400. doi: 10.1093/plankt/fbw023
- Bucklin, A., Yeh, H. D., Questel, J. M., Richardson, D. E., Reese, B., Copley, N. J., & Wiebe, P. H. (2019). Time-series metabarcoding analysis of zooplankton diversity of the NW Atlantic continental shelf. *ICES Journal of Marine Science* , 76 (4), 1162–1176. doi: 10.1093/icesjms/fsz021
- Buttay, L., Miranda, A., Casas, G., González-Quirós, R., & Nogueira, E. (2015). Long-term and seasonal zooplankton dynamics in the northwest Iberian shelf and its relationship with meteo-climatic and hydrographic variability. *Journal of Plankton Research* , 38 (1), 106–121. doi: 10.1093/plankt/fbv100
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods* , 13 (7), 581–583. doi: 10.1038/nmeth.3869
- Carroll, E. L., Gallego, R., Sewell, M. A., Zeldis, J., Ranjard, L., Ross, H. A., ... Constantine, R. (2019). Multi-locus DNA metabarcoding of zooplankton communities and scat reveal trophic interactions of a generalist predator. *Scientific Reports* , 9 (1), 1–14. doi: 10.1038/s41598-018-36478-x
- Carugati, L., Corinaldesi, C., Dell’Anno, A., & Danovaro, R. (2015). Metagenetic tools for the census of marine meiofaunal biodiversity: An overview. *Marine Genomics* , 24 , 11–20. doi: 10.1016/j.margen.2015.04.010
- Chamberlain, S. A., & Szöcs, E. (2013). taxize: taxonomic search and retrieval in R. *F1000Research* , 2 , 191. doi: 10.12688/f1000research.2-191.v1
- Chao, A. (1984). Nonparametric Estimation of the Number of Classes in a Population. *Scandinavian Journal of Statistics* , 11 (4), 265–270. doi: 10.1214/aoms/1177729949
- Clarke, K. R., & Warwick, R. M. (2001). A further biodiversity index applicable to species lists: Variation in taxonomic distinctness. *Marine Ecology Progress Series* , 216 , 265–278. doi: 10.3354/meps216265
- Clarke, L. J., Soubrier, J., Weyrich, L. S., & Cooper, A. (2014). Environmental metabarcodes for insects: In silico PCR reveals potential for taxonomic bias. *Molecular Ecology Resources* , 14 (6), 1160–1170. doi: 10.1111/1755-0998.12265
- Corell, J., & Rodríguez-Ezpeleta, N. (2013). Tuning of protocols and marker selection to evaluate the diversity of zooplankton using metabarcoding. *Revista de Investigaciones Marinas* , 21 (2), 19–39. Retrieved from http://www.azti.es/rim/wp-content/uploads/2014/05/Revista-Marina-21_2.pdf
- Creer, S., Deiner, K., Frey, S., Porazinska, D., Taberlet, P., Thomas, W. K., ... Bik, H. M. (2016). The ecologist’s field guide to sequence-based identification of biodiversity. *Methods in Ecology and Evolution* , 7 (9), 1008–1018. doi: 10.1111/2041-210X.12574
- Cristescu, M. E. (2014). From barcoding single individuals to metabarcoding biological communities: Towards an integrative approach to the study of global biodiversity. *Trends in Ecology and Evolution* , 29 (10), 566–571. doi: 10.1016/j.tree.2014.08.001
- Damien, P., Pasqueron de Fommervault, O., Sheinbaum, J., Jouanno, J., Camacho-Ibar, V. F., & Duteil, O. (2018). Partitioning of the Open Waters of the Gulf of Mexico Based on the Seasonal and Interannual Variability of Chlorophyll Concentration. *Journal of Geophysical Research: Oceans* , 123 (4), 2592–2614. doi: 10.1002/2017JC013456
- Deagle, B. E., Jarman, S. N., Coissac, E., Pompanon, F., & Taberlet, P. (2014). DNA metabarcoding and the

- cytochrome c oxidase subunit I marker: Not a perfect match. *Biology Letters* , 10 (9), 2–5. doi: 10.1098/rsbl.2014.0562
- Djurhuus, A., Pitz, K., Sawaya, N. A., Rojas-Márquez, J., Michaud, B., Montes, E., ... Breitbart, M. (2018). Evaluation of marine zooplankton community structure through environmental DNA metabarcoding. *Limnology and Oceanography: Methods* , 16 (4), 209–221. doi: 10.1002/lom3.10237
- Elliott, D. T., Pierson, J. J., & Roman, M. R. (2012). Relationship between environmental conditions and zooplankton community structure during summer hypoxia in the northern Gulf of Mexico. *Journal of Plankton Research* , 34 (7), 602–613. doi: 10.1093/plankt/fbs029
- Espinasse, B., Carloti, F., Zhou, M., & Devenon, J. L. (2014). Defining zooplankton habitats in the gulf of lion (NW Mediterranean Sea) using size structure and environmental conditions. *Marine Ecology Progress Series* , 506 , 31–46. doi: 10.3354/meps10803
- Everaert, G., Deschutter, Y., De Troch, M., Janssen, C. R., & De Schamphelaere, K. (2018). Multimodel inference to quantify the relative importance of abiotic factors in the population dynamics of marine zooplankton. *Journal of Marine Systems* , 181 (February), 91–98. doi: 10.1016/j.jmarsys.2018.02.009
- Färber Lorda, J., Athié, G., Camacho Ibar, V., Daessle, L. W., & Molina, O. (2019). The relationship between zooplankton distribution and hydrography in oceanic waters of the Southern Gulf of Mexico. *Journal of Marine Systems* , 192 (March 2018), 28–41. doi: 10.1016/j.jmarsys.2018.12.009
- Gasca, R., Suárez-Morales, E., & Haddock, S. H. D. (2007). Symbiotic associations between crustaceans and gelatinous zooplankton in deep and surface waters off California. *Marine Biology* , 151 (1), 233–242. doi: 10.1007/s00227-006-0478-y
- Gaston, K. J. (2009). Geographic range limits of species. *Proceedings of the Royal Society B: Biological Sciences* , 276 (1661), 1391–1393. doi: 10.1098/rspb.2009.0100
- Gazonato Neto, A. J., Silva, L. C. da, Saggio, A. A., & Rocha, O. (2014). Zooplankton communities as eutrophication bioindicators in tropical reservoirs. *Biota Neotropica* , 14 (4). doi: 10.1590/1676-06032014001814
- Gluchowska, M., Trudnowska, E., Goszczko, I., Kubiszyn, A. M., Blachowiak-Samolyk, K., Walczowski, W., & Kwasniewski, S. (2017). Variations in the structural and functional diversity of zooplankton over vertical and horizontal environmental gradients en route to the Arctic Ocean through the Fram Strait. *PLoS ONE* , 1–26. doi: 10.1371/journal.pone.0171715
- Hamilton, P. (2007). Eddy statistics from Lagrangian drifters and hydrography for the northern Gulf of Mexico slope. *Journal of Geophysical Research: Oceans* , 112 (9), 1–16. doi: 10.1029/2006JC003988
- Hawkins, B. A., & Diniz-filho, J. A. F. (2004). ‘Latitude’and geographic patterns in species richness. *Ecography* , 27 (2), 268–272.
- Heberle, H., Meirelles, V. G., da Silva, F. R., Telles, G. P., & Minghim, R. (2015). InteractiVenn: A web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinformatics* , 16 (1), 1–7. doi: 10.1186/s12859-015-0611-3
- Heimeier, D., Lavery, S., & Sewell, M. A. (2010). Using DNA barcoding and phylogenetics to identify Antarctic invertebrate larvae: Lessons from a large scale study. *Marine Genomics* , 3 (3–4), 165–177. doi: 10.1016/j.margen.2010.09.004
- Helenius, L. K., Leskinen, E., Lehtonen, H., & Nurminen, L. (2017). Spatial patterns of littoral zooplankton assemblages along a salinity gradient in a brackish sea: A functional diversity perspective. *Estuarine, Coastal and Shelf Science* , 198 , 400–412. doi: 10.1016/j.ecss.2016.08.031
- Hereu, C. M., Arteaga, M. C., Galindo-Sánchez, C. E., Herzka, S. Z., Batta-Lona, P. G., & Jiménez-Rosenberg, S. P. A. (2020). Zooplankton summer composition in the southern Gulf of Mexico with emphasis

- on salp and hyperiid amphipod assemblages. *Journal of the Marine Biological Association of the United Kingdom* , 1–16. doi: 10.1017/s0025315420000715
- Howson, U. A., Buchanan, G. A., & Nickels, J. A. (2017). Zooplankton Community Dynamics in a Western Mid-Atlantic Lagoonal Estuary. *Journal of Coastal Research* , 78 , 141–168. doi: 10.2112/si78-012.1
- Jennifer E., P., & Mary N., A. (2001). Interactions of pelagic cnidarians and ctenophores with fish: a review. *Hydrobiologia* , 451 , 27–44.
- Johnston, E. L., Mayer-Pinto, M., & Crowe, T. P. (2015). Chemical contaminant effects on marine ecosystem functioning. *Journal of Applied Ecology* , 52 (1), 140–149. doi: 10.1111/1365-2664.12355
- Jouanno, J., Ochoa, J., Pallàs-Sanz, E., Sheinbaum, J., Andrade-Canto, F., Candela, J., & Molines, J. M. (2016). Loop current frontal eddies: Formation along the campeche bank and impact of coastally trapped waves. *Journal of Physical Oceanography* , 46 (11), 3339–3363. doi: 10.1175/JPO-D-16-0052.1
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution* , 30 (4), 772–780. doi: 10.1093/molbev/mst010
- Larke, L. J., Beard, J. M., Swadling, K. M., & Deagle, B. E. (2017). Effect of marker choice and thermal cycling protocol on zooplankton DNA metabarcoding studies. *Ecology and Evolution* , 7 (3), 873–883. doi: 10.1002/ece3.2667
- Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., . . . Machida, R. J. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. *Frontiers in Zoology* , 10 (1), 1–14. doi: 10.1186/1742-9994-10-34
- Linacre, L., Lara-Lara, R., Camacho-Ibar, V., Herguera, J. C., Bazán-Guzmán, C., & Ferreira-Bartrina, V. (2015). Distribution pattern of picoplankton carbon biomass linked to mesoscale dynamics in the southern gulf of Mexico during winter conditions. *Deep-Sea Research Part I: Oceanographic Research Papers* , 106 , 55–67. doi: 10.1016/j.dsr.2015.09.009
- Lindeque, P. K., Parry, H. E., Harmer, R. A., Somerfield, P. J., & Atkinson, A. (2013). Next generation sequencing reveals the hidden diversity of zooplankton assemblages. *PLoS ONE* , 8 (11), 1–14. doi: 10.1371/journal.pone.0081327
- Machida, R. J., Leray, M., Ho, S.-L., & Knowlton, N. (2017). Data Descriptor: Metazoan mitochondrial gene sequence reference datasets for taxonomic assignment of environmental samples. *Scientific Data* , 4 (September 2016), 1–7. doi: 10.1038/sdata.2017.27
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal* , 17 (1), 10–12. Retrieved from <http://dx.doi.org/10.14806/ej.17.1.200>
- McArdle, B. H., & Anderson, M. J. (2001). Fitting Multivariate Models To Community Data : *Ecology* , 82 (1), 290–297.
- Muller-Karger, F. E., Smith, J. P., Werner, S., Chen, R., Roffer, M., Liu, Y., . . . Enfield, D. B. (2015). Natural variability of surface oceanographic conditions in the offshore Gulf of Mexico. *Progress in Oceanography* , 134 , 54–76. doi: 10.1016/j.pocean.2014.12.007
- Palumbi, S. R., Sandifer, P. A., Allan, J. D., Beck, M. W., Fautin, D. G., Fogarty, M. J., . . . Wall, D. H. (2009). Managing for ocean biodiversity to sustain marine ecosystem services. *Frontiers in Ecology and the Environment* , 7 (4), 204–211. doi: 10.1890/070135
- Parmar, T. K., Rawtani, D., & Agrawal, Y. K. (2016). Bioindicators: the natural indicator of environmental pollution. *Frontiers in Life Science* , 9 (2), 110–118. doi: 10.1080/21553769.2016.1162753

- Perez-Brunius, P., Garcia-Carrillo, P., Dubranna, J., Sheinbaum, J., & Candela, J. (2013). Direct observations of the upper layer circulation in the southern Gulf of Mexico. *Deep-Sea Research Part II: Topical Studies in Oceanography* , 85 , 182–194. doi: 10.1016/j.dsr2.2012.07.020
- Pinol, J., Mir, G., Gomez-Polo, P., & Agusti, N. (2015). Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Molecular Ecology Resources* , 15 (4), 819–830. doi: 10.1111/1755-0998.12355
- Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLoS ONE* , 5 (3). doi: 10.1371/journal.pone.0009490
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., ... Glockner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research* , 41 (D1), 590–596. doi: 10.1093/nar/gks1219
- Ratnasingham, S., & Hebert, P. D. N. (2007). BOLD : The Barcode of Life Data System (www.barcodinglife.org). *Molecular Ecology Notes* , 7 , 355–364. doi: 10.1111/j.1471-8286.2006.01678.x
- Salmeron-Garcia, O., Zavala-Hidalgo, J., Mateos-Jasso, A., & Romero-Centeno, R. (2011). Regionalization of the gulf of Mexico from space-time chlorophyll-a concentration variability. *Ocean Dynamics* , 61 (4), 439–448. doi: 10.1007/s10236-010-0368-1
- Sheinbaum, J., Athie, G., Candela, J., Ochoa, J., & Romero-Arteaga, A. (2016). Structure and variability of the Yucatan and loop currents along the slope and shelf break of the Yucatan channel and Campeche bank. *Dynamics of Atmospheres and Oceans* , 76 , 217–239. doi: 10.1016/j.dynatmoce.2016.08.001
- Stefanni, S., Stanković, D., Borme, D., de Olazabal, A., Juretić, T., Pallavicini, A., & Tirelli, V. (2018). Multi-marker metabarcoding approach to study mesozooplankton at basin scale. *Scientific Reports* , 8 (1), 1–13. doi: 10.1038/s41598-018-30157-7
- Steinberg, D. K., & Landry, M. R. (2017). Zooplankton and the Ocean Carbon Cycle. *Annual Review of Marine Science* , 9 (1), 413–444. doi: 10.1146/annurev-marine-010814-015924
- Trudnowska, E., Gluchowska, M., Beszczynska-Möller, A., Blachowiak-Samolyk, K., & Kwasniewski, S. (2016). Plankton patchiness in the Polar Front region of the west Spitsbergen Shelf. *Marine Ecology Progress Series* , 560 (November), 1–18. doi: 10.3354/meps11925
- Uribe-Martínez, A., Aguirre-Gómez, R., Zavala-Hidalgo, J., Ressler, R., & Cuevas, E. (2019). Unidades oceanográficas del Golfo de México y áreas adyacentes : La integración mensual de las características biofísicas superficiales Oceanographic units of Gulf of Mexico and adjacent areas : The monthly integration of surface biophysical features. *Geofísica Internacional* , 58 (4), 295–315.
- Usov, N., Khaitov, V., Smirnov, V., & Sukhotin, A. (2019). Spatial and temporal variation of hydrological characteristics and zooplankton community composition influenced by freshwater runoff in the shallow Pechora Sea. *Polar Biology* , 42 (9), 1647–1665. doi: 10.1007/s00300-018-2407-1
- Vause, B. J., Morley, S. A., Fonseca, V. G., Jazdzewska, A., Ashton, G. V., Barnes, D. K. A., ... Peck, L. S. (2019). Spatial and temporal dynamics of Antarctic shallow soft-bottom benthic communities: Ecological drivers under climate change. *BMC Ecology* , 19 (1), 1–14. doi: 10.1186/s12898-019-0244-x
- Vazquez-Baeza, Y., Pirrung, M., Gonzalez, A., & Knight, R. (2013). EMPERor : a tool for visualizing high-throughput microbial community data. *Gigascience* , 2 (Nov), 1–4.
- Wang, Q., Garrity, G. M., Tiedje, J. M., & Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology* , 73 (16), 5261–5267. doi: 10.1128/AEM.00062-07
- WoRMS Editorial Board. (2017). Retrieved from <https://doi.org/10.14284/170>

Xiong, W., Ni, P., Chen, Y., Gao, Y., Li, S., & Zhan, A. (2019). Biological consequences of environmental pollution in running water ecosystems: A case study in zooplankton. *Environmental Pollution* ,252 , 1483–1490. doi: 10.1016/j.envpol.2019.06.055

Yang, J., & Zhang, X. (2019). eDNA metabarcoding in zooplankton improves the ecological status assessment of aquatic ecosystems. *Environment International* , 134 (September 2019), 105230. doi: 10.1016/j.envint.2019.105230

Zhan, A., Bailey, S. A., Heath, D. D., & Macisaac, H. J. (2014). Performance comparison of genetic markers for high-throughput sequencing-based biodiversity assessment in complex communities. *Molecular Ecology Resources* , 14 (5), 1049–1059. doi: 10.1111/1755-0998.12254

Zhang, G. K., Chain, F. J. J., Abbott, C. L., & Cristescu, M. E. (2018). Metabarcoding using multiplexed markers increases species detection in complex zooplankton communities. *Evolutionary Applications* , 11 (10), 1901–1914. doi: 10.1111/eva.12694

Data Accessibility

BioProject ID PRJNA630297

Author Contributions

Conceptualization: FC, MCA, CGS; Data curation: FC, ASF, JRF; Formal analysis: FC, RG, MM, CMH; Funding acquisition: MCA, CGS; Investigation: FC, MCA, CGS; Methodology: ASF, JRF; Project administration: MCA, CGS; Resources: MCA, CGS; Software: FC, RG, MM; Supervision: SH, MCA, CGS; Validation: FC, MCA, SH, MM, CMH, SJR, ASF, JRF, RG, PGBL, CGS; Visualization: FC, MCA, SH, MM, CMH, SJR, ASF, JRF, RG, PGBL, CGS; Writing – original draft: FC; Writing – review & editing: FC, MCA, SH, MM, CMH, SJR, ASF, JRF, RG, PGBL, CGS.

Tables

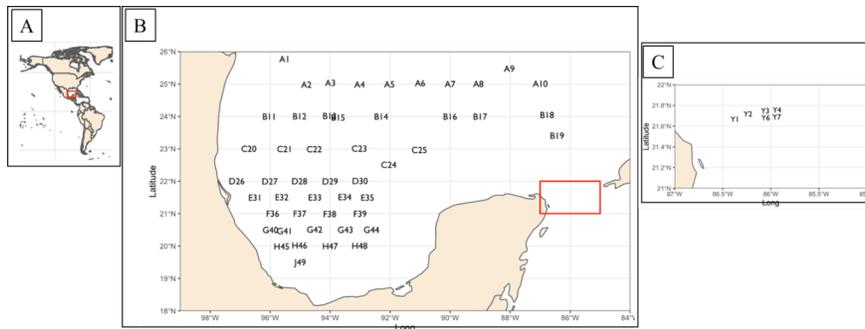
[A]			
	MARGINAL TEST [18S rRNA]	MARGINAL TEST [18S rRNA]	
Variable	Sum squares	Sum squares	Ps
Oxygen	4316.6		10.
Temperature	4260		10.
Longitude	3297.3		7.7
Salinity	3005.5		7.0
Latitude	2241.2		5.1
Density	2087.8		4.8
Fluorescence	1349		3.0
Depth	1130.9		2.5
[C]			
	SEQUENTIAL TESTS [18S rRNA]	SEQUENTIAL TESTS [18S rRNA]	
Variable	AICc	Sum squares	Ps
Oxygen	737.46	4316.6	10.
Longitude	731.03	3364.3	8.6
Temperature	727.82	1989.5	5.2
Salinity	727.47	910.93	2.4

Table 1. Distance-based linear modeling (DistLM) results showing the relationship between environmental variables and stations. Marginal tests show the influence of each parameter in isolation with 18S rRNA [A] and cytochrome oxidase c subunit I (COI)[B]. Results of the sequential tests show the effect of environmental parameters on zooplankton assemblages in the combined model (stepwise selection with adjusted AICc) for

18S rRNA[C] and COI [D] .

[B]			
	MARGINAL TEST [COI]	MARGINAL TEST [COI]	MARGINAL TEST [COI]
Variable	Sum squares	Sum squares	Pseudo-F
Temperature	8063.3	8063.3	7.7494
Density	7996.1	7996.1	7.6806
Oxygen	7508.7	7508.7	7.1838
Longitude	5519.7	5519.7	5.1971
Salinity	4897.7	4897.7	4.5886
Latitude	4456.4	4456.4	4.1607
Depth	2179.7	2179.7	1.999
Fluorescence	1929.4	1929.4	1.766

[D]			
	SEQUENTIAL TESTS [COI]	SEQUENTIAL TESTS [COI]	SEQUENTIAL TESTS [COI]
Variable	AICc	Sum squares	Pseudo-F
Temperature	835.78	8063.3	7.7494
Longitude	833.34	4567.1	4.5202
Oxygen	832.15	3232.6	3.2612



Figures

Figure 1 .[A] : Map of North and South America with a red box enclosing the sampling area within the Gulf of Mexico shown in[B] . [B] Locations of the sampling stations (A1 to J49) within the Gulf of Mexico. Red box delimiting the sampling area around the Yucatan peninsula shown in [C] .[C] Location of the sampling stations (Y1 to Y7) around the Yucatan peninsula.

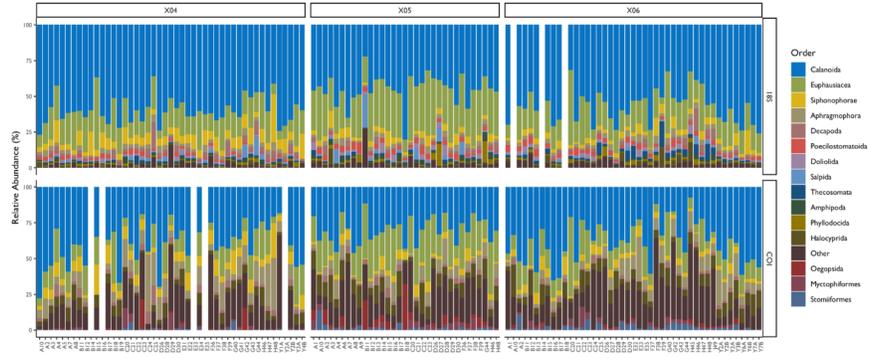


Figure 2: Bar-plot of the major zooplankton taxa at the level of order among all stations from each cruise (XIXIMI-04, XIXIMI-05, and XIXIMI-06) detected for 18S rRNA and cytochrome oxidase c subunit I (COI).

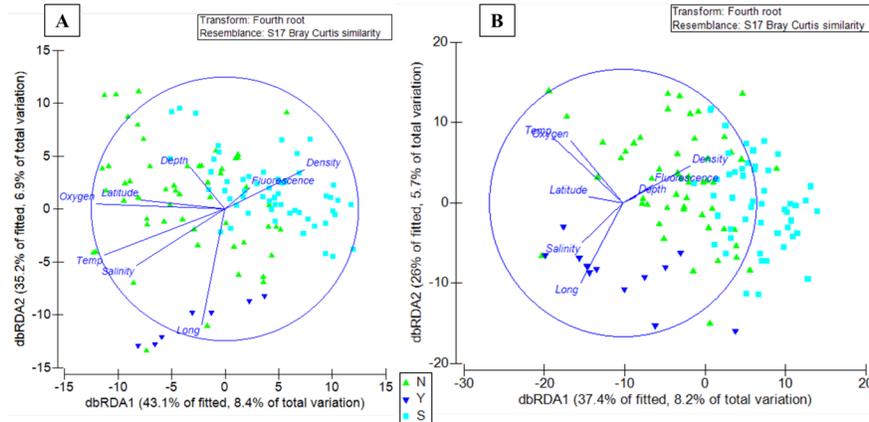


Figure 3: Distance-based redundancy (dbRDA) plots illustrating distance-based linear modeling (DistLM) of the similarities among stations and their relationships with environmental variables for 18S rRNA [A] and cytochrome oxidase c subunit I (COI) [B]. The axis legends include the percent (%) of variation explained by the fitted model and the % of total variation explained by the axis. Stations are highlighted according to the proposed partitioning of the Gulf of Mexico (N = northern, green triangles; S = southern, blue squares; Y = Yucatan peninsula, blue triangles).

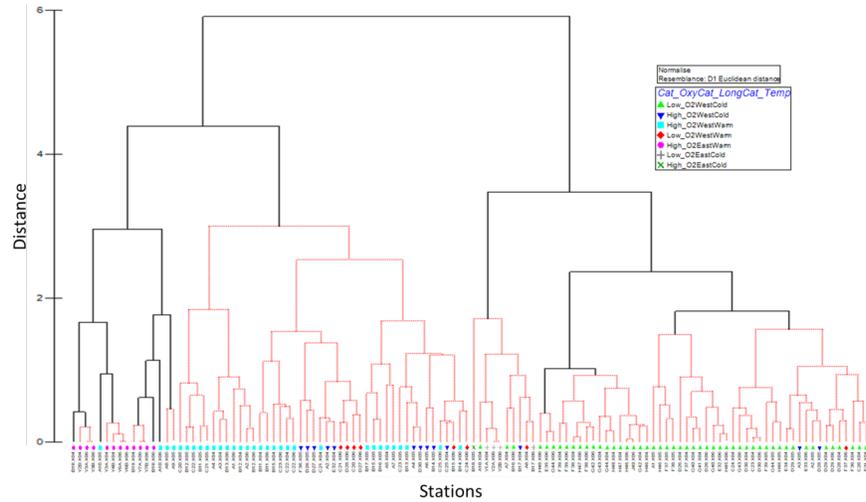


Figure 4: Cluster plot representing the segregation of stations based on oxygen, temperature, and longitude gradients. The stations are represented by individual IDs composed of the sampling line (from A to J, see experimental design in the main text), a sequential numerical identifier, and the sampling date indicated by X4 (XIXIMI-04), X5 (XIXIMI-05), and X6 (XIXIMI-06).

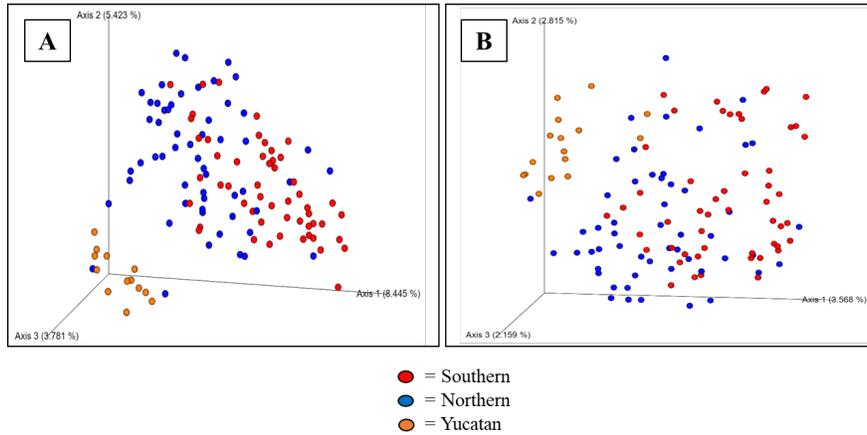


Figure 5: Spatial principal coordinate analysis (PCoA) based on the unweighted Unifrac distance of zooplankton amplicon sequence variants (ASVs) species among the northern, southern, and Yucatan ecoregions detected with 18S rRNA[A] and cytochrome oxidase c subunit I (COI)[B] .

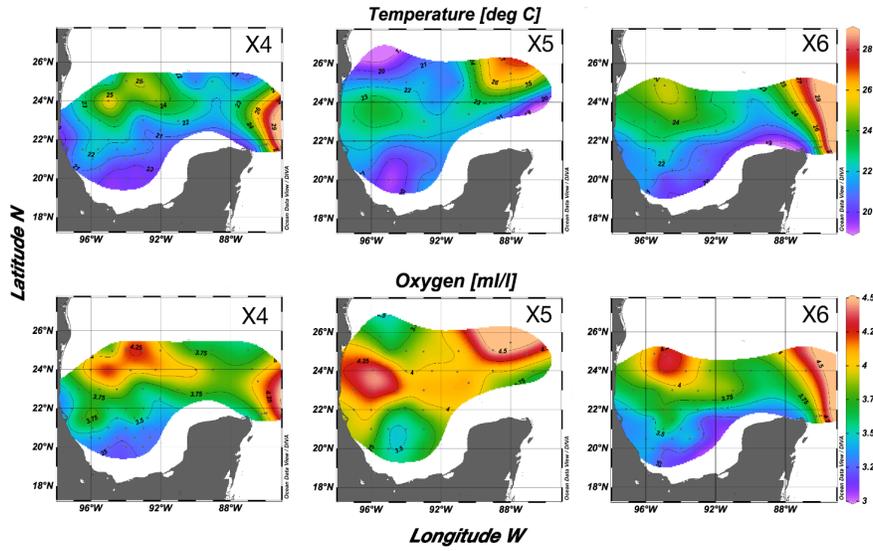


Figure 6: Seasonal variation in temperature and oxygen concentrations during each field expedition for 2015 (XIXIMI-04), (XIXIMI-05), and 2017 (XIXIMI-06). The color axes indicate temperature ($^{\circ}\text{C}$) or oxygen (ml/l) values.

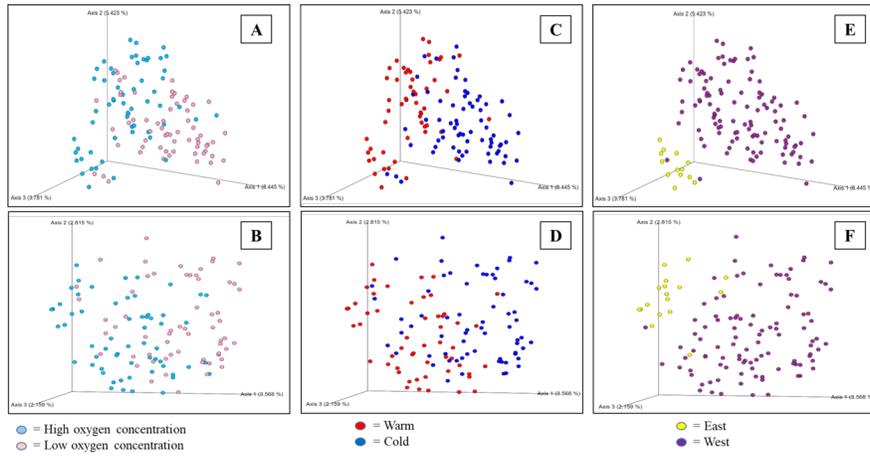


Figure 7: Principal coordinate analysis (PCoA) based on the Jaccard distance of amplicon sequence variants (ASVs) among oxygen, temperature, and longitude discrete variables for 18S rRNA in [A], [C], and [E] and cytochrome oxidase c subunit I (COI) in [B], [D], and [F], respectively.

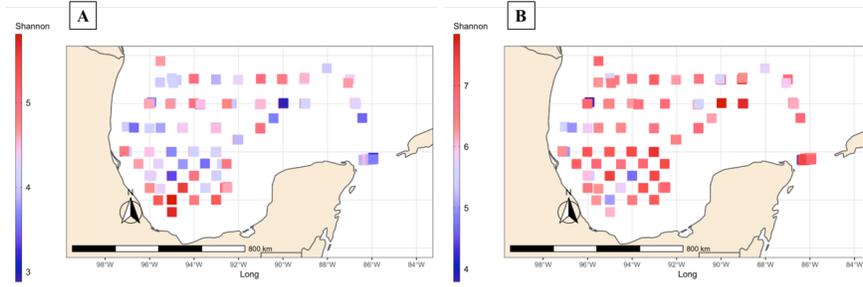


Figure 8: Heat map based on the values of Shannon diversity index from each sampling station detected with 18S rRNA [A] and cytochrome oxidase c subunit I (COI) [B] .

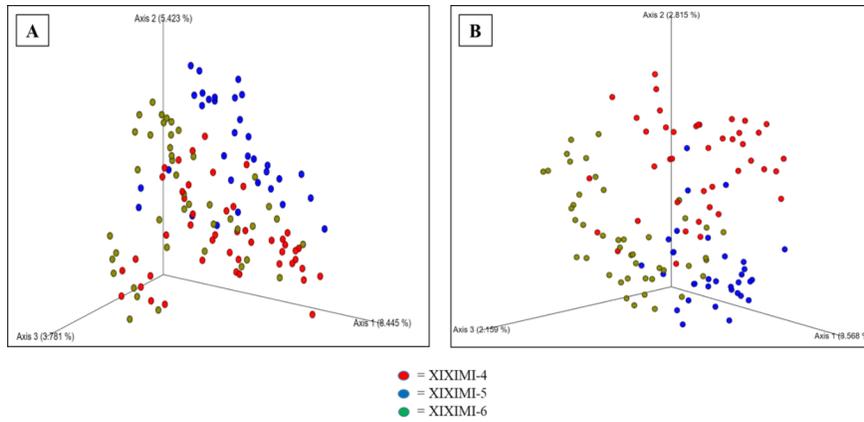


Figure 9: Temporal principal coordinate analysis based on the Jaccard distance of zooplankton amplicon sequence variants (ASV) species among field expeditions detected with 18S rRNA [A] and cytochrome oxidase c subunit I (COI) [B].