Saudi Journal of Biological Sciences 29 (2022) 103342

Contents lists available at ScienceDirect

Saudi Journal of Biological Sciences

journal homepage: www.sciencedirect.com

Original article

Exploring the taxonomic and functional diversity of marine *benthic* micro-Eukaryotes along the Red Sea coast of Jeddah city

Samah S. Abuzahrah [a,*], Mohammed N. Baeshen [a], Ali Alkaladi [a], Noor M. Bataweel [b], Ahmed M. Alhejen [b], Hayam Abdelkader [a,c]

[a] *Department of Biological Sciences, College of Science, University of Jeddah, Saudi Arabia*
[b] *king Fahd medical research center, king Abdul-Aziz university P.O.BOX 80203 Jeddah 21589, Saudi Arabia*
[c] *Virus Research Department, Molecular biology lab., PPRI, ARC., Egypt*

ARTICLE INFO

Article history:
Received 15 December 2021
Revised 7 May 2022
Accepted 10 June 2022
Available online 26 June 2022

Keywords:
Marine metagenomics
18S rDNA Micro-eukaryotic organisms
Red Sea coast
Benthic sedimentary soil
Gene profile
Illumina MiSeq platform

ABSTRACT

Backgrounds: Diverse marine habitats along Jeddah's Red Sea coast support rich biodiversity. Few studies have been done on its diverse communities, especially its microbial counterparts. Metagenomic analysis of marine benthic micro-eukaryotic communities was performed for the first time on the Red Sea coast of Jeddah. This research looks into their community structure and metabolic potential.

Methods: Next-generation sequencing was used to examine the micro-eukaryotic communities of seven sedimentary soil samples from four Jeddah coast locations. After isolating DNA from seven benthic sedimentary soil samples, the 18S rDNA V4 regions were amplified and sequenced on the Illumina MiSeq. It was also verified using an Agilent Technologies 2100 Bioanalyzer with a DNA 1000 chip (Agilent Technologies, Fisher Scientific). A standard curve of fluorescence readings generated by qPCR quantification using the Illumina library was achieved using the GS FLX library. Metagenomic data analysis was used to evaluate the microbial communities' biochemical and enzymatic allocations in studied samples.

Results: Blast analysis showed that the top ten phyla were Annelida, Eukaryota, Diatomea, Porifera, Phragmoplastophyta, Arthropoda, Dinoflagellata, Xenacoelomorpha Nematoda, and uncultured. Annelida was also found in the highest percentage (93%), in the sample M followed by Porifera (64%), the most abundant in the control sample then Eukaryotes (61%), Phragmatoplastophyta (55%), Arthropoda, and Diatomea (the least common) (32%). community diversity analysis: using Shannon and inverse Simpson indices showed sediment composition to be effective. Also, PICRUST2 indicated that the most abundant pathways were pyruvate fermentation to isobutanol, pyrimidine deoxyribonucleotide phosphorylation, adenosine ribonucleotide de novo biosynthesis, guanosine ribonucleotide de novo biosynthesis, NAD salvage pathway I, the super pathway of glyoxylate bypass and aerobic respiration I (cytochrome *c*).

Conclusion: Results showed that high throughput metagenomics could reveal species diversity and estimate gene profiles. Environmental factors appear to be more important than geographic variation in determining the structure of these microbial communities. This study provides the first report of marine benthic micro-eukaryotic communities found on the Red Sea coast of Jeddah and will serve as a good platform for future research.

* Corresponding author at: Department of Biological Sciences, College of Science, University of Jeddah, 21959 Saudi Arabia.
E-mail addresses: sabozahera@uj.edu.sa, dr.samahsulimanw@gmail.com (S.S. Abuzahrah), mnbaeshen@uj.edu.sa (M.N. Baeshen), oalmaghrabi@uj.edu.sa (A. Alkaladi), BataweelN@kau.edu.sa (N.M. Bataweel), aalhejin@kau.edu.sa (A.M. Alhejen), hsabdelkader@uj.edu.sa, hsaabdelkader@arc.sci.eg (H. Abdelkader).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.sjbs.2022.103342
1319-562X/

1. Introduction

About two-thirds of the Earth's surface (~71%) is covered with marine biomass (Bar-On and Milo, 2019). The marine environment is generally divided into pelagic and bottom (or *benthic*) water zones (Samah sulaiman et al., 2021). In spatial coverage, marine sedimentary habitats specify the most significant single ecosystem on the Earth. A diverse range of abundant and ubiquitous meiobenthos inhabiting these sediments is essential globally because they contribute to human well-being and the survival of life on Earth. They mediate vital ecosystem processes such as sediment reworking, nutrient recycling, food web dynamics, pollutant degradation and distribution, and organic matter management (decomposition, mineralization, burial, and storage) (Adão and Meiofauna, 2021).

Red Sea (Al-Bahr Al-Ahmar) represents one of the world's warmest, saltiest, and least explored oligotrophic marine environments, in which extensive diversity of the most productive coral reefs are hosted (Behzad et al., 2016). It is a singular, distinctive ecosystem with an abundance of unique features, representing ecologically one of the most critical marine depots and highly endemic regions in the universe, with high biodiversity (Fine et al., 2019). Therefore, it is considered a region of increasing interest for scientists interested in different research domains (SULAIMAN et al., xxxx). However, among aquatic habitats, its microbiome is one of the least investigated (Behzad et al., 2016). The Red Sea's semi-enclosed form makes it an ideal laboratory for studying the effects of environmental gradients on microbial populations (Pearman et al., 2017).

Marine Benthos is the primary and important component of marine sedimentary biodiversities, which are often employed as markers of environmental change. Depending on their size, benthos includes three classes; microfauna, meiofauna, and macrofauna (Bhadury et al., 2020; Udalov et al., 2021).

Marine microbial *eukaryotes* are crucial components of Plankton and the Benthos ecosystems in all ocean biomes**.** They are, along with cyanobacteria, accoaccount approximately half of the global primary productivity in the ocean (Obiol et al., 2020). They play crucial roles in the operation of marine food webs and biogeochemical processes such as carbon export to the deep ocean and nutrient remineralization (Liu et al., 2021). Many *benthic* taxa are accessible to detection (either individually or collectively as a community) of time-integrated reactions to changes in environmental circumstances because of their largely sessile life, high species richness, and sensitivity to environmental or human influences (Zeppilli et al., 2015).

Metagenomics is a branch of biotechnology that focuses on the analysis of large datasets of biological data. Without using conventional microbiota culture methods, it can reveal the proper microbial composition and explore the genetically rich resources of uncultured microbiota (Ahmad et al., 2021). Metagenomics has advanced the field of marine microbial ecology by revealing previously unknown microbes in various aquatic environments (Grossart et al., 2020). Huge amounts of data created help discover metagenomes, community relationships, adaptive processes, and future applications. It finds new organisms and bioindicators and builds phylogenetic trees (Samah sulaiman et al., 2021). Microbial biogeography and ecosystem functioning, stability, and succession are revealed by functional trait-based metagenomics (Langille et al., 2013). Metagenomics may help researchers better understand microbial communities' biochemical and enzymatic allocations, as well as their interactions with the environment (Ahmad et al., 2019). It is also a robust technology that can help accelerate basic research by investigating microbiology principles and building innovative biotechnological processes and products. Metagenomics finds new biocatalysts in environmental materials (SULAIMAN et al., xxxx). The 18S rRNA gene is the best available global marker for *Eukaryotes* due to its presence in all eukaryotes, its inclusion in universal reference databases, and the availability of universal primers (Hadziavdic et al., 2014). Operational taxonomic units (OTUs) are used to identify taxonomic microorganisms, describe community structures, and evaluate the extent of the scarce biospheres (Adamo et al., 2020).

Next-generation sequencing (NGS), also known as short reading sequencing (50–400 bp), relies on the Illumina platform, which provides hundreds of times faster readings with a 0.1 percent error rate. It's a massively parallel sequencing technology with high output. It allows for the simultaneous sequencing of millions of DNA molecules and, making it the most popular technology for studying microbial diversity using bioinformatic analysis (Kulski et al., 2016). These sequences allow accurate identification of microbial taxa, including non-cultivable and rare taxa. In addition, they provide a comprehensive inventory of all microbial operons and genes present or expressed under various environmental conditions (Mayo et al., 2014). Aiming to make metagenomics more accessible, NGS allows for targeted metagenomics (fixed amplified regions of genomic DNA; as 18S amplicon sequencing) (Amrane and Lagier et al., 2018).

The current study used the 18S metagenomic sequencing library preparation guide to build libraries targeting the variable V4 regions of the 18S rDNA gene. The 18S rRNA gene is the best available global marker for *Eukaryotes* due to its presence in all eukaryotes, its inclusion in universal reference databases, and the availability of universal primers (Hadziavdic et al., 2014). 18S metagenomics application in BaseSpace® analytic environment was used to process the results of paired-end sequencing on the MiSeq system. Using the MiSeq system, the 18S metagenomics study identified complex micro-eukaryotic populations at the species level of the marine samples of the Red Sea coast from Jeddah city. The workflow for 18S metagenomics includes DNA isolation, library preparation, sequencing, and push-button analysis. Using Illumina library preparation, MiSeq technology, and easy analytic software, micro-eukaryotic marine samples from the Red Sea coast- Jeddah city were studied and identified. The microbial communities' biochemical and enzymatic allocations, as well as their interactions with the environment, were evaluated in the studied marine samples. A wide marine area that has been impacted by human activities was chosen for this study because it was not previously included in previous research. It was also decided to keep the contour area out of the way of these activities.

Marine benthic microeukaryotes, more specifically protozoans and micromatazoans communities, of the Northern Red Sea coastal areas of Jeddah City, were the primary focus of this study, which included a metagenomic analysis of total genomic DNA extracted from benthic soil samples from this area and a ribosomal DNA gene study (18 s rDNA). Marine benthic proto-zoans and micromatazoans found on the surface of sedimented soils at various depths and locations were studied to determine their taxonomic diversity. There were other considerations in this study, such as the possible ecological roles that each location may play in its ecosystem and the distribution of each organism vs. depth. Benthic microeukaryotic communities and their molecular diversity should be studied for their functional capabilities.

2. Materials and methods

2.1. Sampling locations

On October 12, 2019, seven sea soil samples were taken from the La Playa beach zone (21.72217339.076362) in Jeddah's north seacoast (Abhur Alshamaliah) (Fig. 1). Each site got two repli-

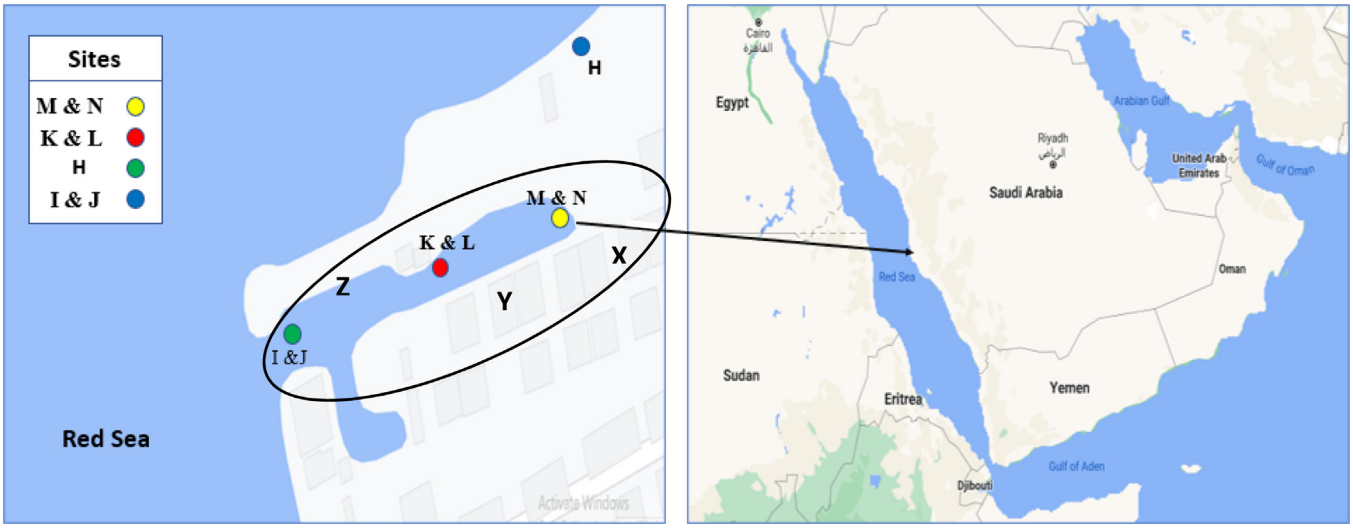

Fig. 1. Jeddah districts map showing the La Playa beach zone locations in the northern seacoast of Jeddah and the locations from which samples were collected in this study. Site (**X**): wet sand samples (**M & N**); site (**Y**): the samples (**K & L**) were taken at ~ 50-meter distance behind the reef towards the shoreline; site (**Z**): samples (**I & J**) obtained 90 m behind the coral reef; control sample (**H**) was obtained from the sandy beach.

cas randomly. The first site (**X**): wet sand samples (**M & N**) from the beach or shore (21.7229417, 39.0785304). The second position (**Y**): where the samples (**K & L**) were taken at ~ 50-meter distance behind the reef towards the shoreline in the shallow water at a depth of ~ 80 cm (21.7221773,350768463). The third site (**Z**): samples (**I & J**) obtained 90 m behind the coral reef at 90 cm depth (21.7220543,39.0763283). A solitary control sample (**H**) was obtained from the sandy beach with no seawater. The samples were kept refrigerated at 4 °C in sterile 50 mL Falcon tubes.

2.2. DNA extraction and quantifications

Total DNA was extracted using the DNeasy PowerSoil Kit (Cat No./ID: 12888–50), following the manufacturer's instructions. The quantity and purity of the eluted DNA were determined by using Nano-Drop 2000 spectrophotometer-Thermo scientific™ (thermo-fisher scientific, DE, USA). DNA purity was evaluated by calculating the absorbance ratio at an optical density (OD) of 260–280 nm. A ratio between 1.7 and 2.0 was required for a suitable polymerase chain reaction (PCR) and indicated the good purity of DNA. The measurements were blanked first to a value of 0–0.2; by pipetting 2 µl of elution buffer used for the DNA extraction. Then, 2.0 µl of each DNA was pipetted and loaded on a clean pedestal. For both blank and sample estimation, the arm was lowered down, and the values of OD at 260–280 nm were recorded (Lear et al., 2018).

2.3. Gel electrophoresis

To confirm the quality and quantity of the isolated DNA of studied samples gel electrophoresis was applied (Lee et al., 2012). Measuring the condition of the extracted DNA was done by running the product on 1% (w/v) agarose gel; prepared by dissolving 1 g of high melting point agarose powder (Agarose low EEO, CSL-AG500, Cleaver Scientific Ltd) with 100 mL 1x TAE buffer (TAE is Tris-Acetate-EDTA) in a microwavable flask. Heat the prepared gel for 1–3 min in the microwave, then cool to 60 °C. Add 5 µl ethidium bromide (EtBr) to the melted agarose gel, then poured in the tray after placing the comb and allowed it to solidify at room temperature for 20–30 mins. The prepared gel was placed into the gel box of the electrophoresis unit, and 1x TAE buffer was added until the gel was covered with the buffer. In the loading step, 5 µl of DNA sample and 2 of bromophenol blue (loading dye) were mixed by pipetting then loaded into the slots for each DNA sample. 5 µl of 1 Kb (250–10,000 bp) DNA Ladder RTU (Ready-to-Use) was added into the first wells as well. Electrophoresis was running out at 120 V for 40 min with 1x TAE as a running buffer. Total DNA was visualized under ultra-violet (UV) light, using a gel documentation system (Picó et al., 2012).

2.4. PCR, and Illumina MiSeq platform

The primers TAReuk454 FWD1 and TAReukREV3 were used for DNA amplification by PCR. The primers were selected to target the variable regions (V4) in the 18S rRNA marker gene. The primers sequences were as follows: TAReuk454 FWD1 (5′-CCAGCAS CYGCGGTAATTCC-3′) and TAReuk REV3 (5′-ACTTTCGTTCTTGA TYRA-3′). The thermocycler conditions were performed as described by BioLabs protocol for Q5 High-Fidelity 2x Master Mix. Step1 was activation of Q5 at 98 °C for 2 min, steps 2, 3, and 4 were repeated ten times: denaturation at 98 °C for 10 s, annealing at 53 °C for 30 s, and extension at 72 °C for 30 s. Then perform the following 15 cycles: denaturation at 98 °C for 10 s, annealing at 48 °C for 30 s, and extension at 72 °C for 30 s—final extension stage of 2 min at 72 °C (Lear et al., 2018). Raw data from Illumina MiSeq of PCR amplicons were processed commercially at Nova Lifetech Pte Ltd (2017128822), Singapore. The 18S metagenomics application was used in the BaseSpace® analytic environment to evaluate the data.

2.5. Data analysis

QIIME (quantitative insights into microbial ecology) 1.9.1v pipeline was used to analyze paired-end read sequences produced by Illumina MiSeq (Caporaso et al., 2011). Quality was evaluated and cut on all readings (maximum homopolymers = 6; minimum length = 200 bp). Following initial filtering, OTUs were grouped using Research (https://www.drive5.com/usearch/) with a threshold value of 97% similarity. BLAST was used to search each read against the SILVA 108 database (https://arb-silva.de/, Nov 2015). Furthermore, singleton reads were eliminated to reduce error. Also, the QIIME software handbook was used to examine both alpha and beta diversity. For the examination of diversity, standard statistical approaches such as the Shannon diversity index, OTUs clustering,

and Chao1 were utilized (Caporaso et al., 2011). Shannon diversity index (H) assesses species richness and evenness, while Chao1 estimates species richness based on abundance (Kim et al., 2017). Rarefaction curves (observed species richness) were used to compare OTU richness between samples. For beta diversity analysis, unweighted paired UniFrac distance matrices were used to create a PCoA (Principal Coordinates Analysis) plot and a UPGMA (unweighted pair group method with arithmetic mean) tree. Unclassified *eukaryotes* are *taxon* names that are not well-known enough to be classified, such as 'unclassified *eukaryote* or 'unclassified *eukaryotic taxon* name'. The workflow of Illumina metagenomics amplicon18S rRNA gen bioinformatic data analysis summarized in Fig. 2.

2.6. PICRUSt analysis

PICRUSt2 (*phylogenetic* investigation of communities by reconstruction of unobserved states) **(https://github.com/picrust/picrust2)** calculates the gene families contributed to a metagenome by *eukaryotes* identified by 18S rRNA. The usage of NGS with PICRUSt2 would provide a valuable platform for studying the intricacies of *eukaryote* community formation and function in an environment. It was first used to estimate the eukaryotic functional composition in specific basic settings, such as the animal and the human gut. It is being utilized to examine the functional evaluation of eukaryotic communities in various contexts such as soil, sediments, and wastewater (Fan et al., 2017).

3. Results

3.1. DNA isolation and quantification from soil samples

DNAs were extracted from seven benthic sedimentary soil samples using the DNeasy Power Soil Kit as described under the Materials and Methods. Nanodrop spectrophotometry was used to evaluate the concentration of gDNAs (Fig. 3**A**) and the 260/280 ratio was used to assess their quality. The purity of the extracted DNAs was confirmed by the ratio260/280, which ranged between 2.3 and 2.9. As indicated in Fig. 3**B**, the gel electrophoresis verified the purity and concentrations of the isolated DNA, the DNA concentrations of each sample were calculated relative to the control sample (H) to validate the nanodrop results by Image lab software Bio-Rad Laboratory **(**Fig. 3**C).**

3.2. OTUs and diversity indices

Illumina MiSeq generated a total of 759,344 raw data reads. After trimming and filtering, 447,94 high-quality reads remained, with a range of 49,515–92,803 reads per sample. The variety of alpha and beta was assessed. Alpha diversity reflects species richness within a location, but beta diversity is significant for understanding differences across sites.

Table 1 highlights the relationships, overlaps, and differences between studied soil samples (H, I, J, K, L, M, and N). "K", "M", and "I" had fewer OTUs and lower diversity than other samples, while "J" showed the highest number of OTUs and diversity. All

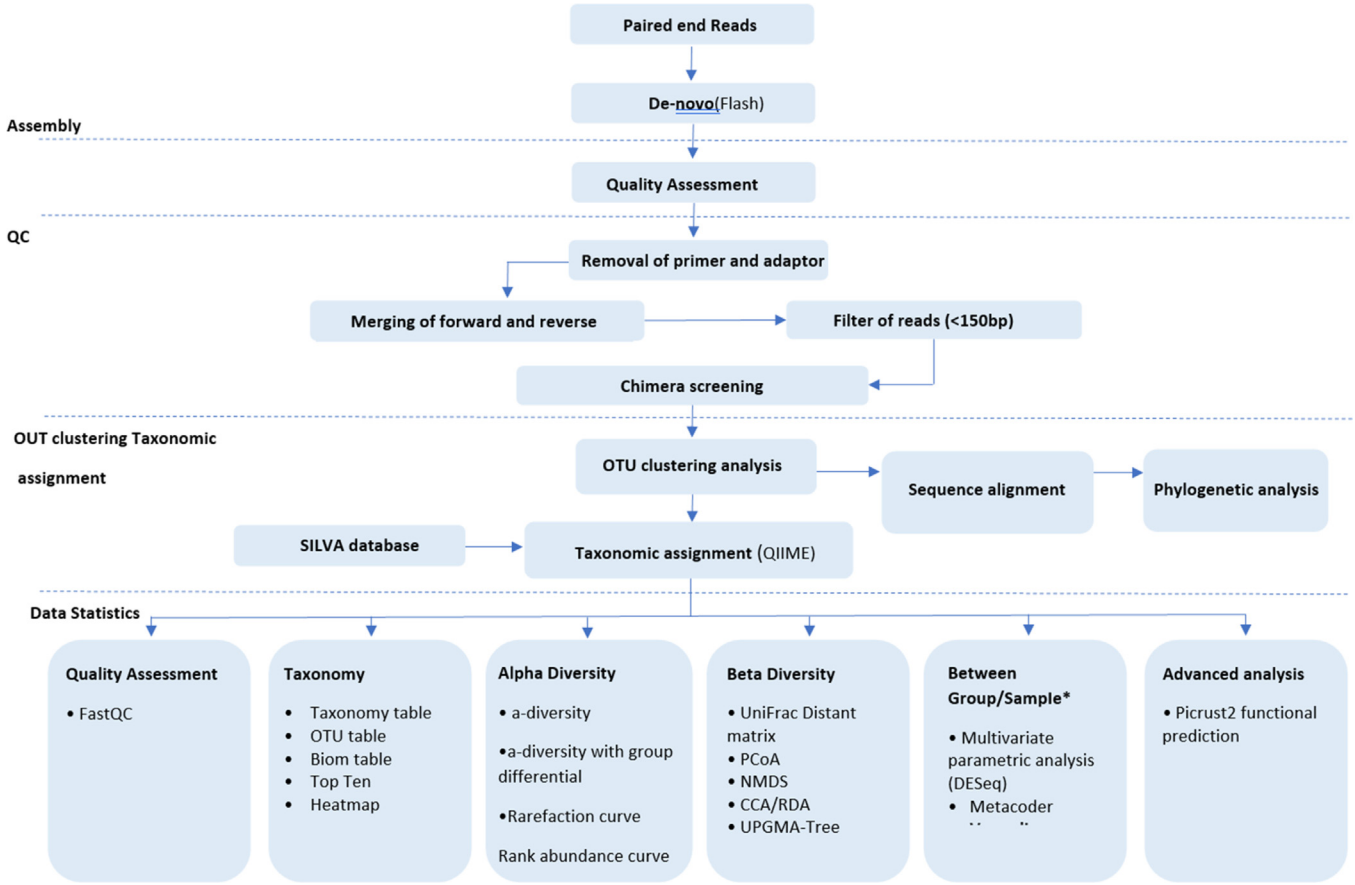

Fig. 2. The workflow of Illumina metagenomics amplicon18S rRNA gen bioinformatic data analysis.

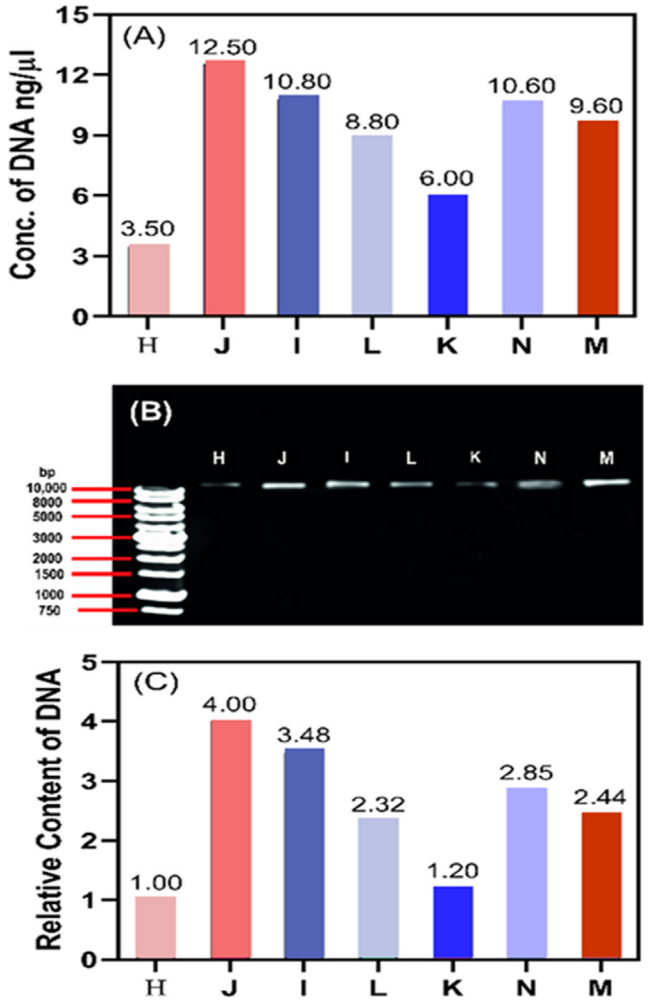

Fig. 3. The quantity and quality of isolated DNA from tested samples by nanodrop (A), gel electrophoresis (B), and the relative color density of DNA bands (C) to control sample (H = 1). Image lab software Bio-Rad Laboratory.

α-diversity indices determined "J" as the highest diversity and richness (Table 1 **&** Fig. 4**A**). However, the ranking for other samples varied across different indices. "K", "M", and "I" had similar OTUs, but showed different results in α-diversity indices. Meanwhile, the ranking of "H", "L", and "N" have also varied across Chao1, ACE, Shannon, Simpson, inverse Simpson, and Fisher indices. While "N" had higher OTUs than "L" in Simpson and inverse Simpson indices, it ranked 3rd in Shannon, Shao, ACE, and Fisher indices. For instance, "J" was the richest sample in all indices except Simpson. The highest OTUs variations across all indices were observed in "J", which ranked 1st in the highest number of different species, richness, concentration, and abundance. The soil *eukaryotic* diversity index did not vary significantly between "I", "K", and "M" (See Table 1).

All rarefaction curves of observed species richness (Fig. 4**A**) have most likely achieved near-saturation. According to the theory that observed species richness is generally connected with total OTUs richness, "J" had the most species richness (781), whereas "K" had the fewest OTUs (349). The rarefaction curves also revealed that "L" had the second greatest OTUs richness, with 619, followed closely by "N", with 545, and then "H", "I", and "M", with 476, 376, and 355, respectively.

The beta diversity of *Benthic* marine *eukaryotes* in coastal regions of La Playa beach zone (Abhur Alshamaliah) in Jeddah's north seacoast was examined using PCoA, which provided the most accurate clusters. The PCoA plot of marine *Benthic eukaryotic* communities (Fig. 4**B**) was utilized to show commonalities and grouping patterns in the data. The PCoA analysis's two axes revealed 21.5 % and 26.1 % of the total variation of the marine *Benthic eukaryotic* communities, respectively. The PCoA figure revealed that the seven sites could be classified into four distinct clusters based on species composition: group (I): ("H", "N", and "M"), group (II): ("J"), group (III): ("L" & "K"), and group (IV): ("I") (Fig. 4**B**).

3.3. Community structure of benthic marine micro-eukaryotes

The current research indicated the richness of *eukaryotic taxonomic* groupings on the Red Sea coast of Jeddah. *Annelida* was the most frequent metazoan *phylae* (93 %), followed by *Porifera* (64%), *Eukaryotes* (61%), *Phragmatoplastophyta* (55%), *Arthropoda*, and *Diatomea* (32%). Fig. 5 and Fig. 6 depict the community architectures of seven samples acquired from distinct locations (community structures of each site: Supplementary S1). The samples "M", "I", "J", "N", and "H" had a higher percentage of *Phylum Annelida* (93%, 57%, 17%, 10%, and 10%, respectively) than the other samples ("K" and "L"), whereas *Porifera* was predominantly present in the control sample and absent from the other samples collected. The *Arthropoda phylum* is the biggest and most successful group in the entire animal kingdom. *Harpacticoida* belonging to order *copepod subphylum Crustacea* was detected only in the control sample "H".

At the *phylum* level, *Annelida* was shown to be the most abundant *phyla* among all seven samples, accounting for 93% in sample "M", 17% in sample "J", 10% in sample "H", and 10% in sample "N". At the same time, *Annelida* was almost absent in "K" & "L".

Porifera (sessile metazoans) was abundant (64% in "H") whereas, *eukaryote* was observed in relatively high numbers (61%) in "N",

Table 1
Alpha diversity indices showing community richness in each sample.

Community richness and diversity							
Sample Name	Observed OTUs	Chao1	ACE	Shannon	Simpson	Inverse Simpson	Fisher
J	781	790.90	789.33	4.45	0.97	31.85	117.37
L	619	630.55	625.99	3.60	0.91	10.53	89.37
N	545	561.92	554.95	3.74	0.94	15.94	77.04
H	467	492.50	488.22	2.26	0.62	2.64	65.82
I	376	384.00	381.06	2.58	0.73	3.66	50.10
M	355	390.68	395.20	0.80	0.22	1.28	46.89
K	349	356.10	353.25	2.94	0.84	6.25	45.98

Chao: species richness estimator estimating the total number of species present in a community by using the frequency of Occurrence of rarer OTUs. If a sample contains many singleton or doubleton, more undetected OTUs likely exist. Moreover, the chao1 index will estimate greater species richness than it would for a sample without rare OTUs. Shanon: A quantitative measure that reflects the number of different types (species) present within a dataset. It also simultaneously considers how evenly the basic entities (individuals) are distributed among types. Inversed Simpson: An indication of how evenly the species is distributed and measure the degree of concentration when individuals are classified into species. ACE: Abundance-Based Coverage estimator.

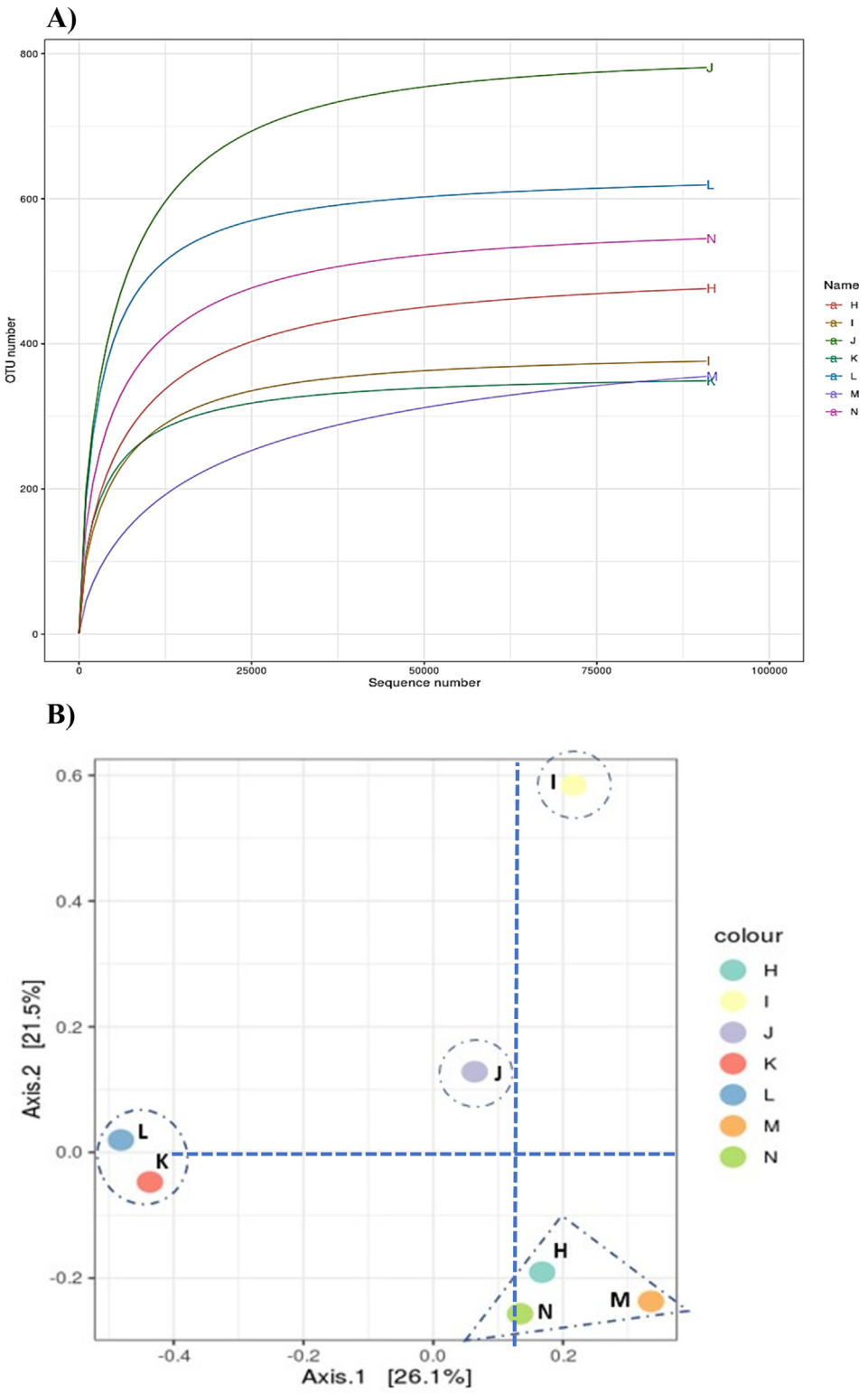

Fig. 4. Rarefaction curves and Beta diversity analysis of PCoA plot of the seven marine *Benthic eukaryotic* samples: **A)** Rarefaction curves showing observed species richness in Red Sea soil samples, the IDs of the samples are on the right side. **B)** Beta diversity analysis of PCoA plot of the seven marine *Benthic eukaryotic* samples. Four groups are formed according to the species composition; group I: (H, M, N), group II: (K, L), group III: I, and Group IV: (J).

60% in "K", 22% in "L", 16% in "J", 12% in "H", and 2% in "M. The specimens (K and L) showed multiple occupations by unclassified *Eukaryotes*; about 15% and 12% fell into the category of "unclassified eukaryotes" (Supplementary S1).

At the class level, based on the total taxonomic abundance in the La Playa beach zone ecosystem, *Polychaeta* were the most prevalent groups discovered, which accounts for 96% population in "M" and 64% in "I" samples. In contrast, 66% *Demospongiae* in

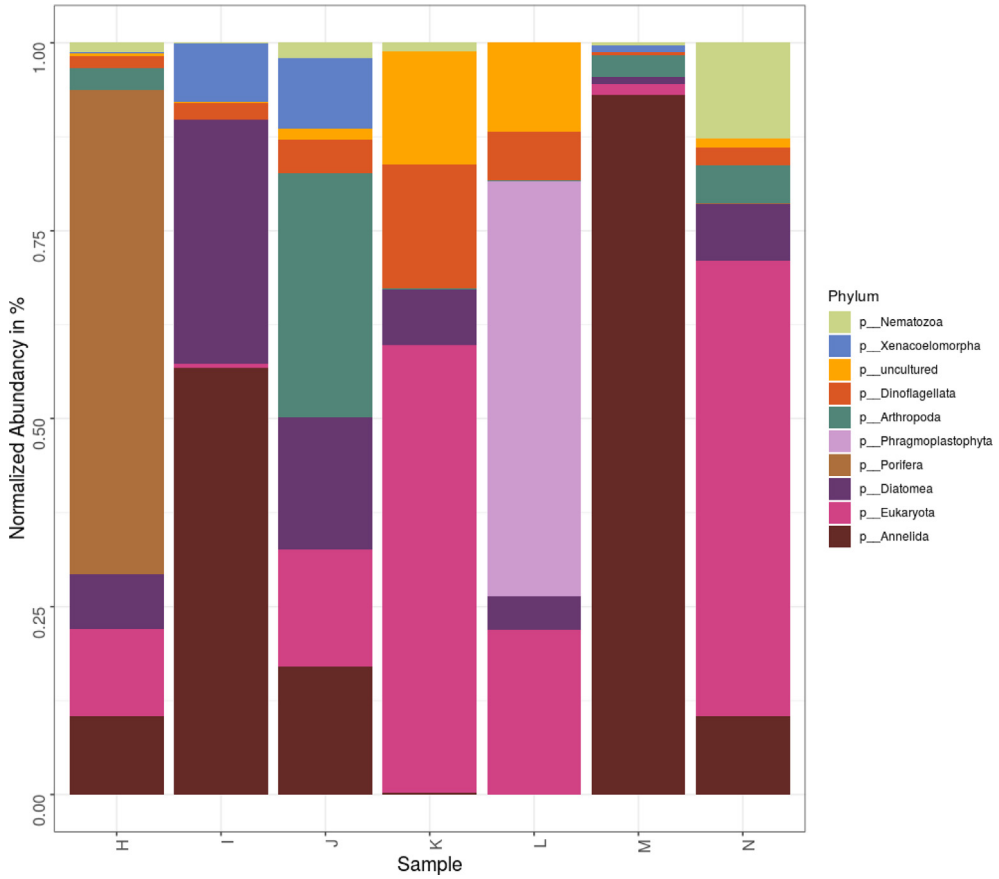

Fig. 5. The abundance rates of dominant phylae in marine *Benthic Eukaryote* communities. The top ten *phylae* are summarized on the right side with the color key.

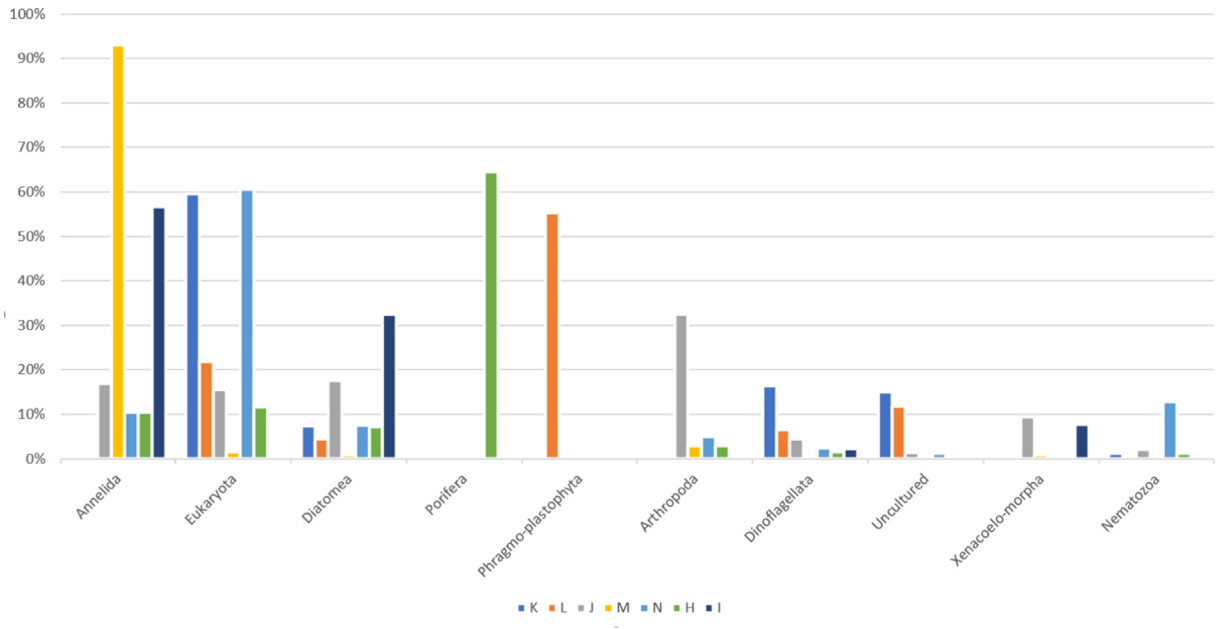

Fig. 6. Taxonomic hierarchy and relative abundance in different soil samples.

"H" and 56% *Embryophyta* were reported in "L". While, 54% *Trebouxiophyceae* in "K", 69% *Ulvophyceae* was recorded in "N" (Supplementary S1).

At the order level, the order *Phyllodocida* was shown to be the most abundant in "M" (96%), "J" (18%), and "N" (10%). Whereas 61% *Demospongiae* dominated in "H", 70% *Polychaeta* in "I", 28% *Harpacticoida* in "J", and *Embryophyta* in "L" (56%). *Trebouxiophyceae* were found to be the most abundant in "K" (55%), "L" (21%), and "N" (5%). *Ulvophyceae* were dominated in "N" (70%), "H" (32%), "J" (14 %), "K" (6%) and " M" (1%) (Supplementary S1).

At the family level, the most abundant family was found to be *Phyllodocida* in "M" (96%), 18% in "J" and 10% in "N". Whereas *Polychaeta* (71%) was dominated in "I". *Ulvophyceae* was dominated in "N" (71%) and 32% in "H". *Embryophyta* were found to be most abundant in "L" (57%), *Trebouxiophyceae* (55%) in "K" and *Harpacticoida* (28%) in "J". Highest similarity between soil samples was observed in *Ulvophyceae* family (Supplementary S1).

The most abundant genus included was *Phyllodocida* (96%) in "M", 19% in "J" and 11% in "N", *Ulvophyceae* (78%) in "N" and 42% in "H". *Protodrilus* was (77%) in "I", *Trebouxiophyceae* (70%) in "K", *Embryophyta* (62%) in "L", and *Ulvophyaceae* (37%) in "H". Important genera may be grouped as follows based on their total abundance (>5%), *Phyllodocida*; 19%, *Acoela*; 14%, and *Navicula*; 12% (Supplementary S1(.

At the species level, the most abundant marine *benthic* micro-*eukaryotic* species indicated in the current study include *Protodrilus gracilis* (segmented worms) (70%), followed by *Symbiochloris sp.* (green algae) (66%), *Phaeophila dendroides* (*Chlorophyta*) (37%), *Thelepinae* sp. (*Annelida*) (36%), *Chromerida* sp. (*Protists*) (32%), *Protodrilus Schaefer* (*Annelida*) (25%), *Oxydromus pugettensi* (*Annelida*) (20%), *Lourinia armada* (*Arthropoda*) (19%), *Terebella* sp. (*Annelida*) (14 %), *Oncholaimidae* sp. (*Nematoda*) (14%), *Odontella aurita* (*Bacil-*

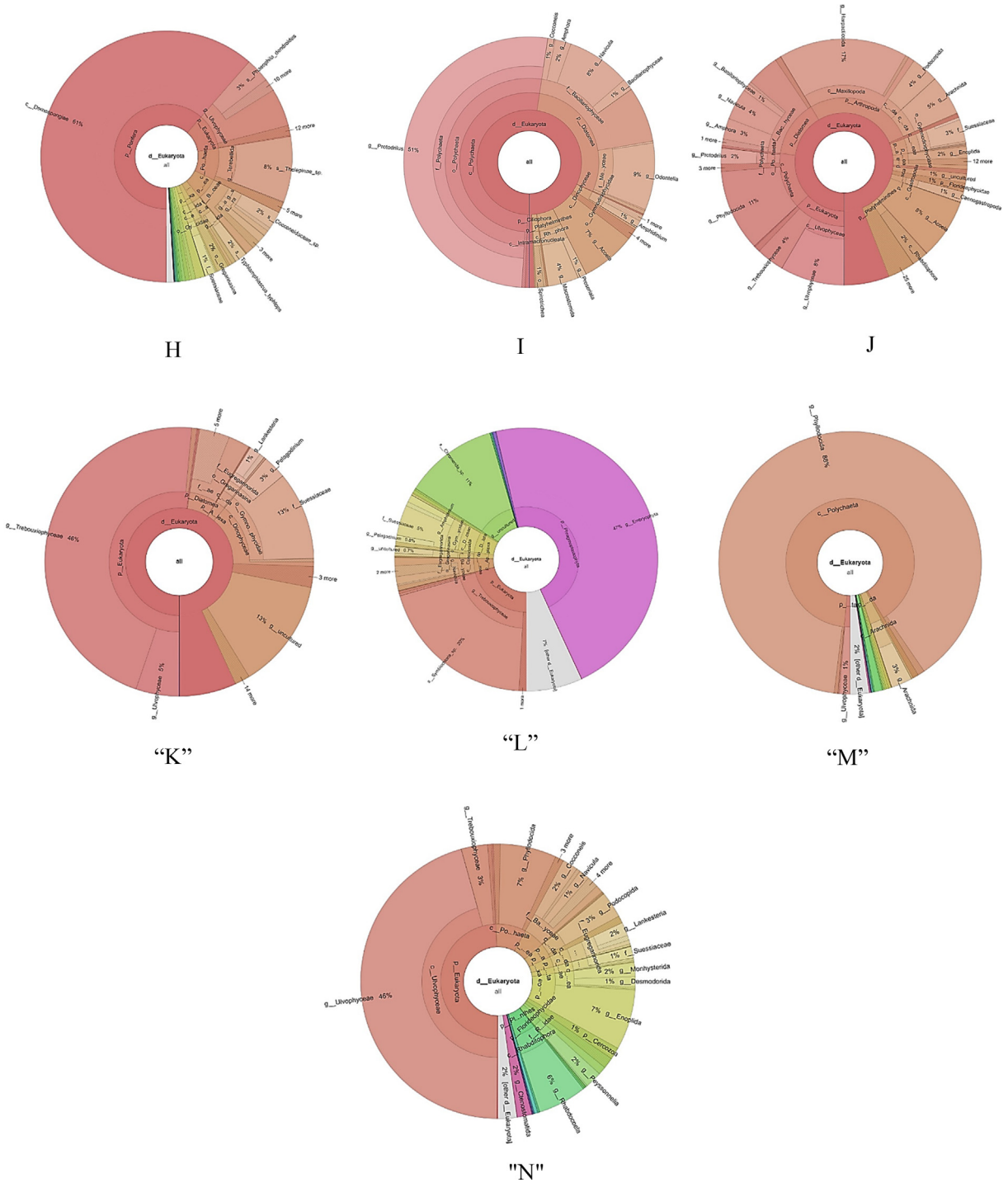

Fig. 7. Krona graph plotted with Krona-Tools-2.7 using taxonomy summary provided by QIIME for taxonomy assignment at the genus level.

Table 2
The percentage of species composition in each soil sample.

Species	Soil samples						
	H	**I**	**J**	**K**	**L**	**M**	**N**
Symbiochloris sp.	4%	0%	7%	**66%**	**56%**	3%	6%
Phaeophila dendroides	**13%**	0%	6%	4%	1%	**12%**	**37%**
Protodrilus gracilis	0%	**70%**	0%	0%	0%	0%	0%
Chromerida sp.	1%	0%	3%	**19%**	**32%**	1%	1%
Thelepinae sp.	**36%**	0%	0%	0%	0%	0%	0%
Protodrilus cf.	0%	0%	4%	0%	0%	**25%**	0%
Cocconeidaceae sp.	**10%**	1%	2%	1%	0%	4%	4%
Oxydromus pugettensi	0%	0%	**20%**	0%	0%	0%	0%
Lourinia armata	0%	0%	**19%**	0%	0%	0%	0%
Terebella sp.	0%	0%	0%	0%	0%	**14%**	0%
Oncholaimidae sp.	0%	0%	0%	0%	0%	0%	**14%**
Odontella aurita	0%	**12%**	1%	0%	0%	0%	0%
Copidognathus sp.	0%	0%	**11%**	0%	0%	0%	1%

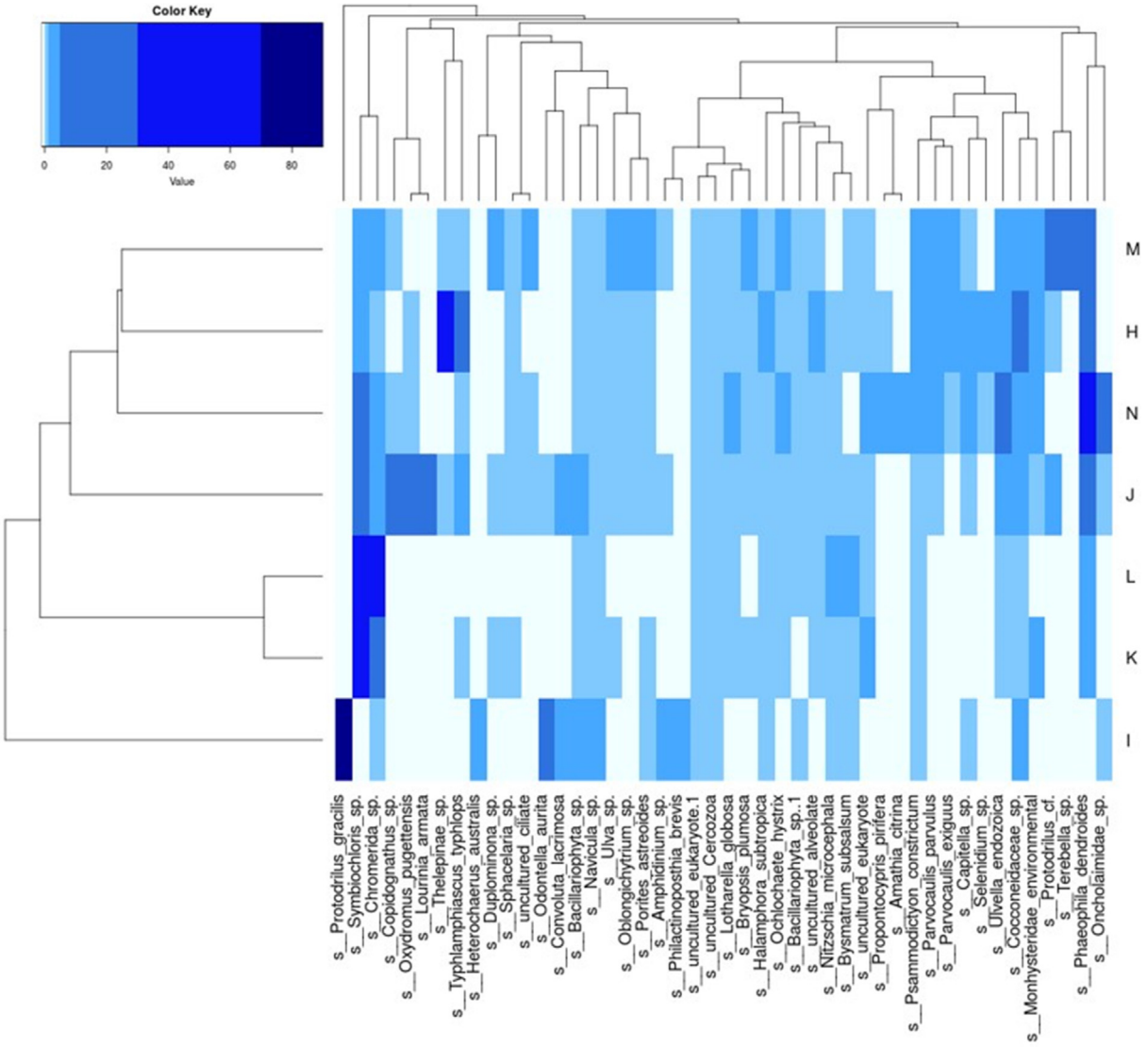

Fig. 8. Species-level heatmap and dendrogram of marine *Benthic eukaryotic* communities with hierarchical two-way cluster analysis. The color intensity in each box shows the percentage of species composition in each sample relative to the color key in the upper left. Soil samples identification is on the right.

lariophyta) (12%), *Copidognathus* sp. (*Arthropoda*) (11%), and *Cocconeidaceae* sp. (A new marine diatom) (10%). Fig. 7 depicts a Krona graph for *taxonomy* assignment at the species level for all samples (See Table 2).

Heat map and dendrogram of marine *benthic* micro-*eukaryotic* species communities based on two-way hierarchical clustering analysis are shown in Fig. 8. Branch lengths in cluster analysis dendrograms correlate to the degree of similarity between branches.

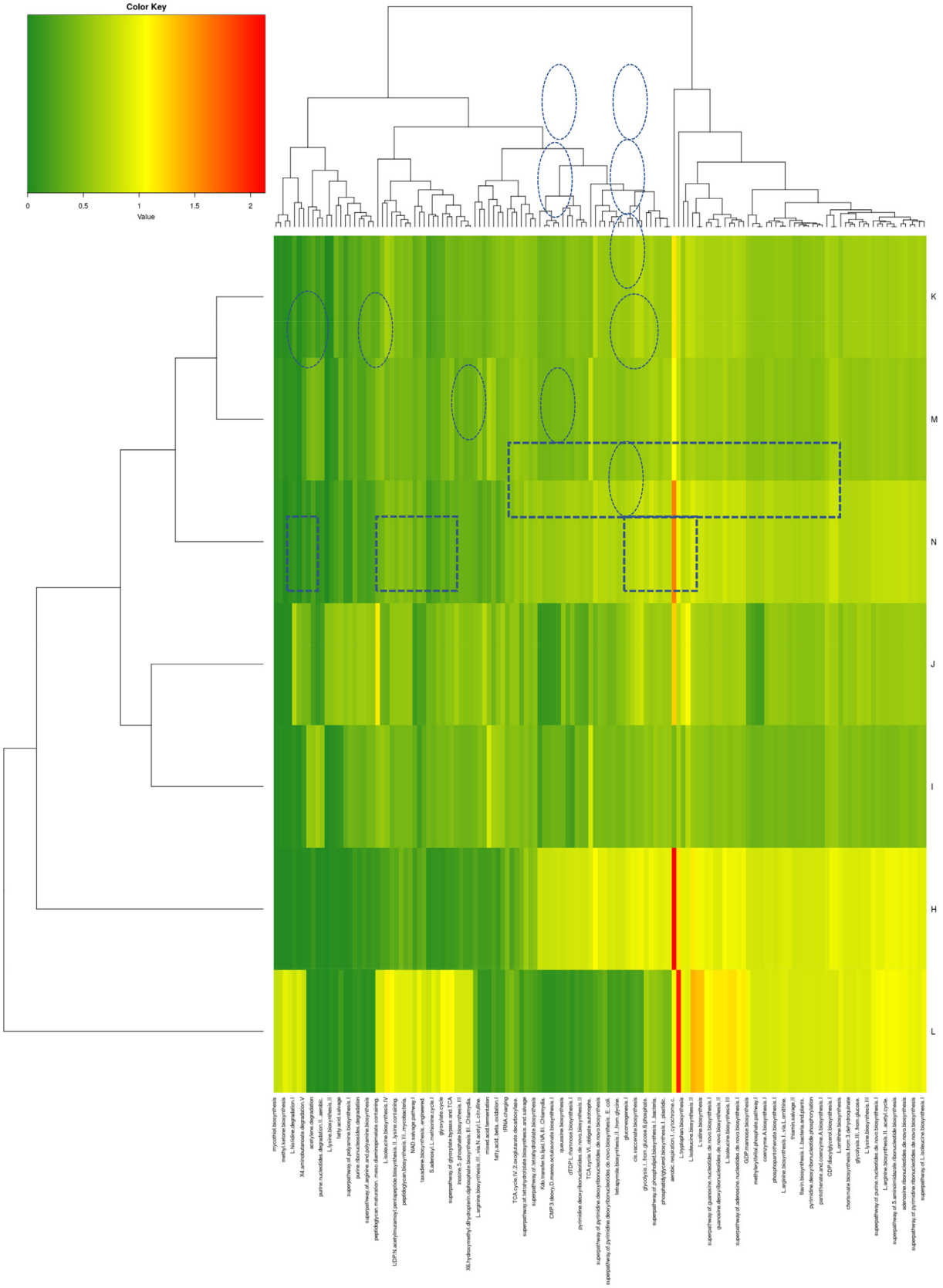

Fig. 9. Picrust2 analysis of predicted functional pathways in marine *benthic* micro-*eukaryotic* communities. Pathways represented on the "X" axis are amino acid metabolism, primary and secondary metabolism, associated metabolites, reactions, enzymes, genes, and energy metabolism. The "Y" axis represents soil samples (I, J, K, L, M, N, and control H). Pathway abundance values (red, orange, green, and yellow) represent the number of genes and are normalized to the total number of genes present in a given pathway from each sample. Samples "H," "N," and "L" showed predicted functional pathways, which are: Pyruvate fermentation to isobutanol, pyrimidine deoxyribonucleotide phosphorylation, adenosine ribonucleotide de novo biosynthesis, guanosine ribonucleotide de novo biosynthesis, NAD salvage pathway I, the super pathway of glyoxylate bypass, L-tryptophan biosynthesis, L-valine biosynthesis, and aerobic respiration I (cytochrome *c*) in a high number. The key color is in the upper left.

Across all samples, the proportion of species composition is shown as a color gradient ranging from deep blue (highest) to light yellow (lowest). The correlation of all species (columns) clustered with soil samples (rows) was shown. All soil samples were highly diverse in terms of species composition. "L" and "K" had a higher degree of similarity (~75%), whereas "M", "H", and "N" were grouped in their clade and had a similarity near to 30%. The sample "J" was assigned to a distinct clade, while "I" was dissimilar to the other clusters (Fig. 9).

The *phylogenetic* investigation of communities by reconstruction of unobserved states (PICRUSt) method was assessed to predict functional profiles using targeted metagenomic data (18S rRNA) gene (Fig. 9**)**. Data used for these predictions included EC-prediction metagenome (enzyme commission numbers) and KO-prediction metagenome (Kyoto ortholog) (Fig. 10**A**) to classify enzymes. KO IDs prediction pathway mapping revealed that sample "L" comprised the highest number of fatty-acyl-CoA synthase, iron complex transport system permease protein, peptide/nickel transport system ATP-binding protein, acetyl-CoA C-acetyltransferase, acetolactate synthase, 3-oxoacyl-[acyl-carrier protein] reductase, and ATP-binding cassette. Whereas alpha-1,3-glucosyltransferase and phosphoenolpyruvate carboxylase were observed in sample "I". Additionally, the remaining samples had varying percentages of these KO pathways Fig. 10**A**.

PICRUSt1 was created in 2013 to predict the functional potential of a bacterial community using marker gene sequencing patterns. PICRUSt2 **(https://github.com/picrust/picrust2)** has been enhanced to include an extended and updated library of gene families and reference genomes, compatibility with any OTU-picking or denoising approach, and phenotypic prediction.

The metagenome's predicted activities revealed that most metacyclic pathways were involved in primary and secondary metabolism, as well as related metabolites, reactions, enzymes, and genes. These hypothesized metacyclic pathways play an essential role in regulating gene expression in biological systems in response to environmental changes, which may help microbial populations adapt in open aquatic habitats.

One of the top 10 metabolic pathways is pyruvate fermentation to isobutanol, this pathway is the manufacture of acetone and butane-1-ol using solvent-producing strains of microorganisms. It was one of the first large-scale industrial fermentation processes to be invented, and during the first half of the 20th century, it ranked second in significance only to ethanol fermentation. Advances in the chemical synthesis of these solvents, along with low oil costs, contributed to the collapse of this commercial fermentation.

The second important pathway is the pyrimidine deoxyribonucleotide phosphorylation; de novo synthesis of ribonucleotides is very expensive, and organisms have adapted to salvage possible precursors from their environment. Since nucleotides cannot be imported into the cell due to the negative charge of the phosphate groups, salvage is confined to free bases and nucleosides.

The adenosine ribonucleotide de novo biosynthesis and guanosine ribonucleotide de novo biosynthesis are very important pathways of nucleotide biosynthesis that are important for nucleic acid synthesis and cell division. The cofactor of energy production, carrier, and vitamin synthesis is NAD salvage pathway I, which plays an important role in cell metabolism.

Super pathway of glyoxylate bypass is a pathway that blends the common prokaryotic TCA cycle with the glyoxylate shunt, a variant that avoids the phases of the TCA cycle in which carbon is lost in the form of CO2. The TCA cycle is a catabolic mechanism that creates energy and reduces power as well as precursors for biosynthesis. The TCA cycle is exceedingly prevalent, and a variant of it exists in virtually all aerobic living beings.

The biosynthesis of very important amino acid pathways includes L-tryptophan biosynthesis and L-valine biosynthesis in the top 10 pathways.

Finally, aerobic respiration I (cytochrome *c*)" pathway gene which responsible for oxidative phosphorylation and energy production in electron transport system was found to be most abundant in samples "H", "N" and "L"." In contrast, nitrate reductase, site-specific DNA-methyltransferase (adenine-specific), Acetolactate synthase, protein-N(pi)-phosphohistidine-sugar phosphotransferase, DNA-directed RNA polymerase, carboxymethylenebutenolidase, acetylornithine deacetylase, monosaccharide-transporting ATPase were highest in sample "I". The most abundant pathways in sample "M" were Alcohol dehydrogenase, DNA-directed DNA polymerase, DNA helicase, RNA helicase, and Iron-chelate-transporting ATPase (Fig. 10**B**). Whereas sample "J" showed considerable occupation by cytochrome-c oxidase, histidine kinase, 3-oxoacyl-[acyl-carrier-protein] reductase, Ribosomal-protein-alanine N-acetyltransferase, glycerol-3-phosphate 1-o-acyltransferase, glucokinase, protein-N(pi)-phosphohistidine-sugar phosphotransferase, DNA-directed RNA polymerase, DNA-directed DNA polymerase, N-acetylmuramoyl-l-alanine amidase, iron-chelate-transporting ATPase, anthranilate synthase, 4-hydroxy-tetrahydrodipicolinate synthase, peptidylprolyl isomerase, acetyl-CoA carboxylase, NADH dehydrogenase, nitrate reductase, thioredoxin-disulfide reductase, and peptide-methionine (S)-S-oxide reductase.

The enzyme commission (EC) prediction function was shown in Fig. 10**C, and A** total of 1846 enzymes have been discovered in the Red Sea soil samples collected in this study. The most abundant predicted enzymes involved in metacyclic pathways in the seven soil samples were:

DNA-directed DNA polymerase, NADH: ubiquinone reductase (H($^{+}$)-translocating), T histidine kinase, serine/threonine (Ser/Thr) kinases, tyrosine (Tyr) kinases, and histidine kinases, peptidylprolyl isomerase, acetolactate synthase, 3-oxoacyl-[acyl-carrier-protein] reductase, glutaminyl-tRNA synthase, RNA helicases, protein-N(pi)-phosphohistidine-sugar phosphotransferase, asparaginyl-tRNA synthase, iron-chelate-transporting ATPase, acetyl-CoA C-acetyltransferase, sarcosine oxidase, N-acetylmuramoyl-l-alanine amidase, ribosomal-protein-and alanine N-acetyltransferase.

4. Discussion

Our current study took a recent trend to discover biodiversity and describe biological communities in modern taxonomic ways that depend on studying the genetic material of living organisms within their environments and classifying them based on the degree of difference and similarity of their DNA away from the old taxonomic methods based on microbial culture, and that what our results showed was similar In the modern methods used, metagenomics and different in the extent of abundance and diversity of living organisms. Also, the lack of studies of the taxonomic structure and functional genomes of micro-eukaryotic organisms along the Red Sea coast increased the value of our current study.

Urbanization, intensive agriculture, digging businesses, and other environmental projects have caused more significant disruptions (Flandroy et al., 2018). Invasive pathogens and metabolic markers have altered microbial community structure and function (Gibbons et al., 2014). Because the Red Sea is heavily used by humans, we need to understand how human behavior affects eukaryotic consortia composition. Metagenomic analysis of Red Sea sampling locations over time may provide a comprehensive picture of the benthic eukaryotic community. Due to the diversity of aquatic microbes, it is difficult to understand how species metabolism and interactions affect microbial community dynamics and

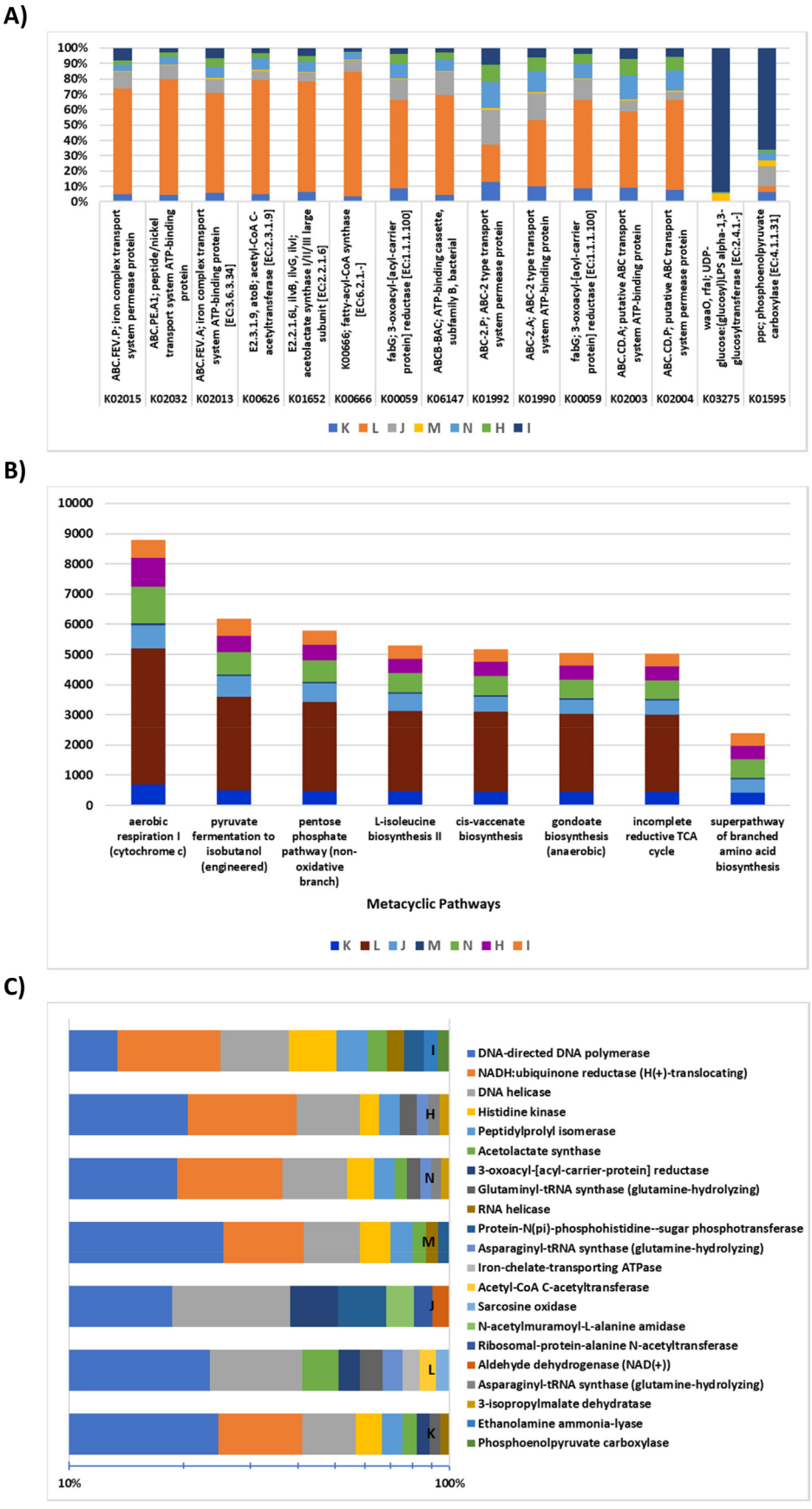

Fig. 10. Metagenomic analysis of of *benthic* micro-eukaryotes in Red Sea soil samples: **A)** KO prediction pathways related to metagenomic sequences of 18S rDNA, **B)** Potential metabolic pathways encoded by the metagenomic sequences (MetaCyc) involved in primary and secondary metabolism, associated metabolites, reactions, enzymes, and genes, and **C)** Prediction of the enzymatic functions of micro-eukaryotic communities using enzymatic function prediction tool (ECPred) (https://github.com/cansyl/ECPred).

biogeochemical changes (Grossart et al., 2019). Using DNA-based methods can help us understand long-term dynamics of marine microeukaryotic communities (Lind and Pollard, 2021). Metagenomic analysis reveals long-lost taxa, OTUs, and functional attributes of micro-eukaryotes associated with Red Sea soil. Cost-effective whole-genome DNA sequencing (WMGS) has recently become possible, opening new potential for applied marine microbiological research. WMGS have improved our understanding of the vast diversity of microorganisms found in aquatic environments (Chandran et al., 2020), allowed more in-depth investigations to identify microbial species taxonomy from environmental samples (Kumar Awasthi et al., 2020).

In line with the (Giner et al., 2020) study, he present the first survey investigating pico-nuclear community composition and phylogenetic activity using Illumina sequencing of 18S rRNA genes amplified from RNA and RNA extracts. In the current study, metagenomic sequences of DNA were used to explore the micro-eukaryotic diversity in seven marine soil samples. The five most abundant taxonomic phyla were *Annelida* (93%), which is the most common *Metazoan*, followed by *Porifera* (64%), *eukaryote* (61%), *Phragmoplastophyta* (55%), and *Arthropoda* (32%). *Annelida* was previously recognized as a typical soil *phylum* (Glasby et al., 2021). Marine sponges (*phylum Porifera*) are ancient *Metazoans* that live in marine and freshwater ecosystems worldwide (Steinert et al., 2019). Previous research has demonstrated that the community structure and diversity of soil micro *Eukaryotes* are dominated by *Metazoa* (28% reads; 17% OTUs), *Cercozoa* (11% reads; 27% OTUs), and *Ciliophora* (4% reads; 8% OTUs) (Mundra et al., 2021). The *phylum Annelida* is split into two classes: *Clitellata* (mostly freshwater creatures) and *Polychaeta* (mostly saltwater animals; marine organisms). There are around 9000 species known worldwide (Glasby et al., 2021).

According to the present research, *Polychaeta* accounted for 93% of marine species. Despite their abundance, *polychaetes* are often injured and killed during the collecting process, making morphological identification difficult. In the current study, three *polychaetes* species, *Thelepinae* sp., *Protodrilus Schaefer,* and *Oxydromus pugettensi,* were identified in the collected soil samples of the Red Sea coast (Martin et al., 2017). Martin et al. (Martin et al., 2017) recorded 199 macrofaunal species of which polychaetes, crustaceans, and *mollusks* are the most important *taxa* along the Red Sea coast of Saudi Arabia (Martin et al., 2017). (Kim et al., 2016) has recorded around 300 polychaete species from Korea. A complete list of *polychaete* species, including 1037 species, 345 *genera*, and 60 families, has been reported for South China and the Philippines (Salazar-Vallejo et al., 2014). Eight hundred thirty-six polychaete species have been recorded from Greece (Faulwetter et al., 2011). The previous report demonstrated that *polychaete* diversity in Canadian waters is underestimated since one-quarter of the 333 morphologically recognized species are composed of two or more putative cryptic species (Carr et al., 2011).

Although the variety of soil protists has long been underestimated, scientific advances such as environmental DNA analysis and high-throughput DNA sequencing can reveal previously undiscovered types (Mahe et al., 2017; de Vargas et al., 2015). Protists comprise the bulk of *Eukaryotes* on a microscopic level, except land plants (*Embryophyta*), mammals, and potentially fungi. *Protists* exhibit an infinite variety of morphological and lifestyle characteristics, ranging from microscopic *picoeukaryotes* taxonomic groups smaller than many bacteria (Moon-van der Staay et al., 2001; Caron et al., 2009), to *plasmodia*-forming slime mold *taxa* and the world's largest single-celled organisms (*Caulerpa*). Other *protists* include 'naked *amoeboid*' and plated forms (e.g., *diatoms* and *Testate amoebae*). They may be *photoautotrophs* (*algae*) or *heterotrophs* (*protozoa*), or *mixotrophs*, accumulating carbon in both ways (Geisen et al., 2017). Numerous *protists* interact mutualistically or parasitically with animals, plants, fungi, and other *protists* (de Vargas et al., 2015). The order *Harpacticoida*, the *phylum Arthropoda*, *subphylum Crustacea* were detected in the current study only in the control sample "H" but not in the other soil samples. *Harpacticoida* is the second biggest meiofaunal group identified in marine sediments after nematodes (Boxshall, 2010).

The current study performed a functional profile based on targeted metagenomic data (18S rDNA) using the PICRUSt2 software tool developed by (Langille et al., 2013). PICRUSt2 builds relationships between taxonomic groups and their associated functions, resulting in more comprehensive and biologically valuable information.

Furthermore, KO IDs were exposed to pathway mapping for micro-*eukaryotic* metabolism in marine soil samples. It was discovered that sample "L" contained the most fatty-acyl-CoA synthase, iron complex transport system permease protein, peptide/nickel transport system ATP-binding protein, acetyl-CoA C-acetyltransferase, acetolactate synthase, 3-oxoacyl-[acyl-carrier protein] reductase, and ATP-binding cassette. These fatty acid transport protein family members perform critical roles in transporting exogenous fatty acids into the cell and intracellular fatty acid homeostasis (Melton et al., 2011). However, alpha-1,3-glucosyltransferase and phosphoenolpyruvate carboxylase were observed in sample "I", which is vital in assimilating atmospheric CO_2 and supporting carbon-nitrogen interactions (Wang et al., 2016).

Numerous common metabolic pathways were identified, demonstrating the diversity and activity of the micro-*eukaryotic* population among Red Sea samples. Pyruvate fermentation to isobutanol (engineered), aerobic respiration I (cytochrome *c*), L-isoleucine biosynthesis II, pentose phosphate pathway, *cis*-vaccinate biosynthesis, and TCA cycle I (prokaryotic) were the most common pathways present in studied samples. All metabolic activities (respiration, fermentation, and photosynthesis) and energy metabolism are dependent on redox reactions between electron donor and acceptor. During redox reactions, microorganisms produce ATP energy for biosynthesis and other cellular functions (Bertrand et al., 2015). There are several oxidative metabolic processes. The primary pathway for carbohydrates is the glycolysis pathway, which is associated with the tricarboxylic acid cycle (TAC). Other metabolic processes can break down glucose, including the pentose phosphate pathways. Pyruvate is a crucial component of the fermentative metabolism of carbohydrates. It is the source of a variety of fermented products produced by various pathways (Sharma et al., 2020).

The quest for enzymes that may be used in green industrial applications is significantly increasing (Renn et al., 2021). The worldwide market for industrial enzymes is predicted to reach $7 billion by 2023, with a 5% annual growth rate (Dewan et al., 2014). The main obstacle in commercial enzymes is that they cannot survive harsh industrial conditions and are not sufficiently stabilized. As a result, the worldwide market requires enzymes that can tolerate a combination of multiple harsh environments (pH, temperature, salinity, organic solvents, and oxidation) in a natural and reproducible manner (Sarmiento et al., 2015; Dumorné et al., 2017; Kara and Liese, 2019).

A possible source of such enzymes is the proteomes of microorganisms that survive in salinity-rich environments such as the Red Sea (Dumorné et al., 2017; Hough et al., 1999). These creatures have developed cellular and molecular mechanisms that allow them to adapt to environmental extremes such as high or low temperatures, acidic or basic pH, high salinity, and high metal concentrations (Eichler, 2001). Because their natural habitats are comparable to those found in industrial processes (Kara and Liese, 2019:), these polyextremophile enzymes have the potential

to satisfy industrial needs (Eichler, 2001; Niehaus et al., 1999; Den Burg et al., 2003; Gomes and Steiner, 2004; Raddadi et al., 2015; Coker et al., 2016).

In this work, the taxonomic analysis of the 18S rRNA sequence of eukaryotic samples from the Red Sea soil sediments enabled us to identify the existent micro-eukaryotic communities and find their metabolic activities. Our results revealed a wide range of enzymes that can be of a potential industrial application, for example, DNA-directed DNA polymerase, NADH: ubiquinone reductase, DNA helicase, Histidine kinase, peptidylprolyl isomerase, Acetolactate synthase, reductase, and others.

DNA-directed DNA polymerase enzymes are utilized in molecular biology for PCR diagnosis (Ishino and Ishino, 2014). Complex I (NADH: ubiquinone oxidoreductase) is the first enzyme in many aerobically respiring organisms' membrane-bound electron transport chain (Hirst et al., 2005; Sazanov et al., 2007). This enzyme is thought to be a substantial source of oxidative stress in cells related to neuromuscular disorders and aging (Esterházy et al., 2008).

The DNA helicases are an essential class of nucleic acid-metabolizing enzymes involved in genome maintenance and cellular homeostasis. DNA helicases control gene expression and chromosomal organization and resolve dynamic DNA structures that cause genomic instability (Brosh et al., 2021). DNA helicase is an essential reagent in pharmaceutical and biotechnological applications.

5. Conclusions

A high throughput metagenomic technology was found to be useful in discovering the diversity of species and estimating functional gene profiles as well as the microbial communities' biochemical and enzymatic allocations in studied samples. The structure of these microbial communities appears to be influenced more by environmental factors than by geographic variations. As a result, the current study suggests that there is room for further investigation of the Red Sea micro-eukaryote for the production of diverse enzymes and secondary metabolites to meet the rising demands of the pharmaceutical and biotechnological industries for the biosynthesis of important biological molecules that could be used in the medical and pharmaceutical fields. For future research, this study's findings provide the first report of marine benthic micro-eukaryotic communities found on the Red Sea coast of Jeddah, Saudi Arabia, and will serve as a good platform for many potentials.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors would like to acknowledge the Department of Biology, Faculty of Science, King Abdulaziz University, and Prof. Sherif Edris Ahmed for their technical support. The authors would also like to express their thanks to La Playa beach, Jeddah for facilitating the samples collection.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.sjbs.2022.103342.

References

Bar-On, Y.M., Milo, R., 2019. The biomass composition of the oceans: a blueprint of our blue planet. Cell 179 (7), 1451–1454.
Sulaiman,, S. et al., 2021. Environmental genomics and biodiversity of macro- and microbenthic communities in the Red Sea coast of Jeddah city. Int. J. Res. Environ. Sci. (IJRES) 7 (1), 38–53.
Adão, H., *Metazoan Meiofauna: Benthic Assemblages for Sustainable Marine and Estuarine Ecosystems*, in *Encyclopedia of the UN Sustainable Development Goals*. 2021, Springer International Publishing. p. 1-22.
Behzad, H. et al., 2016. Metagenomic studies of the Red Sea. Gene 576 (2), 717–723.
Fine, M. et al., 2019. Coral reefs of the Red Sea — challenges and potential solutions. Region. Stud. Marine Sci. 25, 100498.
SULAIMAN, S., et al., *Environmental Genomics and Biodiversity of Macro-and Microbenthic Communities in the Red Sea Coast of Jeddah City.*
Pearman, J.K. et al., 2017. Microbial planktonic communities in the Red Sea: high levels of spatial and temporal variability shaped by nutrient availability and turbulence. Sci. Rep. 7 (1), 6611.
Bhadury, P., et al., *Biodiversity of Benthic Fauna in Chilika Lagoon*, in *Ecology, Conservation, and Restoration of Chilika Lagoon, India*. 2020, Springer International Publishing. p. 365-397.
Udalov, A. et al., 2021. Integrity of benthic assemblages along the arctic estuarine-coastal system. Ecol. Ind. 121, 107115.
Obiol, A. et al., 2020. A metagenomic assessment of microbial eukaryotic diversity in the global ocean. Mol. Ecol. Resour. 20 (3), 718–731.
Liu, X. et al., 2021. Composition change and decreased diversity of microbial eukaryotes in the coastal upwelling waters of South China Sea. Sci. Total Environ. 795, 148892.
Zeppilli, D. et al., 2015. Is the meiofauna a good indicator for climate change and anthropogenic impacts? Mar. Biodivers. 45 (3), 505–535.
Ahmad, T. et al., 2021. Metagenomic analysis exploring taxonomic and functional diversity of bacterial communities of a Himalayan urban fresh water lake. PLoS ONE 16 (3), e0248116.
Grossart, H.P., et al., *Linking metagenomics to aquatic microbial ecology and biogeochemical cycles*. 2020. **65**: p. S2-S20
Langille, M.G.I. et al., 2013. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat. Biotechnol. 31 (9), 814–821.
Ahmad, T. et al., 2019. Metagenomics in the Search for Industrial Enzymes. In: Advances in Enzyme Technology. Elsevier, pp. 419–451.
Hadziavdic, K. et al., 2014. Characterization of the 18S rRNA gene for designing universal eukaryote specific primers. PLoS ONE 9 (2), e87624.
Adamo, M. et al., 2020. Metabarcoding on both environmental DNA and RNA highlights differences between fungal communities sampled in different habitats. PLoS ONE 15 (12), e0244682.
Kulski, J.K., *Next-Generation Sequencing — An Overview of the History, Tools, and "Omic" Applications*, in *Next Generation Sequencing - Advances, Applications and Challenges*. 2016, InTech.
Mayo, B. et al., 2014. Impact of next generation sequencing techniques in food microbiology. Curr. Genom. 15 (4), 293–309.
Amrane, S., Lagier, J.-C., 2018. Metagenomic and clinical microbiology. Hum. Microb. J. 9, 1–6.
Lear, G., et al., Methods for the extraction, storage, amplification and sequencing of DNA from environmental samples. 2018. **42**(1): p. 10-50A
Lee, P.Y., et al., *Agarose gel electrophoresis for the separation of DNA fragments*. 2012 (62): p. e3923
Picó, Y., *Chemical analysis of food: Techniques and applications*. 2012: Academic Press.
Lear, G. et al., 2018. Methods for the extraction, storage, amplification and sequencing of DNA from environmental samples. N. Z. J. Ecol. 42 (1), 10–50A.
Caporaso, J.G., et al., *Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample*. Proceedings of the National Academy of Sciences of the United States of America, 2011. **108 Suppl 1**(Suppl 1): p. 4516-4522.
Kim, B.-R. et al., 2017. Deciphering diversity indices for a better understanding of microbial communities. J. Microbiol. Biotechnol. 27 (12), 2089–2093.
Fan, X.-Y. et al., 2017. Temporal dynamics of bacterial communities and predicted nitrogen metabolism genes in a full-scale wastewater treatment plant. RSC Adv. 7 (89), 56317–56327.
Flandroy, L. et al., 2018. The impact of human activities and lifestyles on the interlinked microbiota and health of humans and of ecosystems. Sci. Total Environ. 627, 1018–1038.
Gibbons, S.M. et al., 2014. Human and environmental impacts on river sediment microbial communities. PLoS ONE 9 (5), e97435.
Grossart, H.P. et al., 2019. Linking metagenomics to aquatic microbial ecology and biogeochemical cycles. Limnol. Oceanogr. 65 (S1).
Lind, A.L. and K.S.J.M. Pollard, *Accurate and sensitive detection of microbial eukaryotes from whole metagenome shotgun sequencing*. 2021. **9**(1): p. 1-18
Chandran, H., M. Meena, and K. Sharma, *Microbial Biodiversity and Bioremediation Assessment Through Omics Approaches*. 2020. **1**(9)
Kumar Awasthi, M. et al., 2020. Metagenomics for taxonomy profiling: tools and approaches. Bioengineered 11 (1), 356–374.
Giner, C.R., et al., Marked changes in diversity and relative activity of picoeukaryotes with depth in the world ocean. 2020. **14**(2): p. 437-449
Glasby, C.J., Erséus, C., Martin, P., 2021. Annelids in extreme aquatic environments: diversity, adaptations and evolution. Diversity 13 (2), 98.
Steinert, G. et al., 2019. *Prokaryotic diversity and community patterns in antarctic continental shelf sponges*. Front. Marine Sci. 6.

Glasby, C.J., Erséus, C., Martin, P., 2021. Annelids in extreme aquatic environments: diversity. Adapt. Evol. 13 (2), 98.
Martin, D. et al., 2017. Another brick in the wall: population dynamics of a symbiotic species of Oxydromus (Annelida, Hesionidae), described as new based on morphometry. J. Contribut. Zool. 86 (3), 181–211.
Kim, H. et al., 2016. Metagenomic analysis of the marine coastal invertebrates of South Korea as assessed by Ilumina MiSeq. Animal Cells Syst. 21 (1), 37–44.
Salazar-Vallejo, S.I. et al., 2014. Polychaete species (Annelida) described from the Philippine and China Seas . Zootaxa 3842 (1), 1.
Faulwetter, S. et al., 2011. An account of the taxonomy and distribution of Syllidae (Annelida, Polychaetes) in the eastern Mediterranean, with notes on the genus Prosphaerosyllis San Martín, 1984 in the Mediterranean. ZooKeys 150, 281–326.
Carr, C.M. et al., 2011. A tri-oceanic perspective: DNA barcoding reveals geographic structure and cryptic diversity in Canadian polychaetes. PLoS ONE 6 (7), e22232.
Mahe, F. et al., 2017. Parasites dominate hyperdiverse soil protist communities in Neotropical rainforests. Nat. Ecol. Evol. 1 (4), 91.
de Vargas, C. et al., 2015. Ocean plankton. Eukaryotic plankton diversity in the sunlit ocean. Science 348 (6237), 1261605.
Moon-van der Staay, S.Y., De Wachter, R., Vaulot, D., 2001. Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. Nature 409 (6820), 607–610.
Caron, D.A. et al., 2009. Defining DNA-based operational taxonomic units for microbial-eukaryote ecology. Appl. Environ. Microbiol. 75 (18), 5797–5808.
Geisen, S. et al., 2017. Soil protistology rebooted: 30 fundamental questions to start with. Soil Biol. Biochem. 111, 94–103.
Boxshall, G., *"Harpacticoida". WoRMS. World Register of Marine Species.* . 2010. **Retrieved May 22, 2010.** .
Melton, E.M. et al., 2011. Human fatty acid transport protein 2a/very long chain acyl-CoA synthetase 1 (FATP2a/Acsvl1) has a preference in mediating the channeling of exogenous n-3 fatty acids into phosphatidylinositol. J. Biol. Chem. 286 (35), 30670–30679.
Wang, N. et al., 2016. Genome-wide analysis of phosphoenolpyruvate carboxylase gene family and their response to abiotic stresses in soybean. Sci. Rep. 6, 38448.
Bertrand, J.-C., et al., *Environmental microbiology: fundamentals and applications*. 2015: Springer.
Sharma, R. et al., 2020. Microbial fermentation and its role in quality improvement of fermented foods. Fermentation 6 (4), 106.
Renn, D., et al., *Novel Enzymes From the Red Sea Brine Pools: Current State and Potential.* 2021. **12**(3267)
Dewan, S.J.B.R., Wellesley, *Global markets for enzymes in industrial applications.* 2014.
Sarmiento, F., Peralta, R., Blamey, J.M., 2015. Cold and hot extremozymes: industrial relevance and current trends. Front. Bioeng. Biotechnol. 3, 148.
Dumorné, K., et al., *Extremozymes: a potential source for industrial applications.* 2017.
Kara, S. and A.J.I.E.A. Liese, *Process considerations for the application of enzymes.* 2019: p. 71-94
Hough, D.W. and M.J.J.C.o.i.c.b. Danson, *Extremozymes.* 1999. **3**(1): p. 39-46.
Eichler, J.J.B.a., *Biotechnological uses of archaeal extremozymes.* 2001. **19**(4): p. 261-278
Niehaus, F., et al., *Extremophiles as a source of novel enzymes for industrial application.* 1999. **51**(6): p. 711-729
Van Den Burg, B.J.C.o.i.m., *Extremophiles as a source for novel enzymes.* 2003. **6**(3): p. 213-218.
Gomes, J., W.J.F.t. Steiner, and Biotechnology, *The biocatalytic potential of extremophiles and extremozymes.* 2004. **42**(4): p. 223-225
Raddadi, N., et al., *Biotechnological applications of extremophiles, extremozymes and extremolytes.* 2015. **99**(19): p. 7907-7913
Coker, J.A.J.F., *Extremophiles and biotechnology: current uses and prospects.* 2016. **5**
Ishino, S., Ishino, Y., 2014. DNA polymerases as useful reagents for biotechnology - the history of developmental research in the field. Front. Microbiol. 5, 465.
Hirst, J., 2005. Energy transduction by respiratory complex I – an evaluation of current knowledge. Biochem. Soc. Trans. 33 (3), 525–529.
Sazanov, L.A., 2007. Respiratory complex I: mechanistic and structural insights provided by the crystal structure of the hydrophilic domain. Biochemistry 46 (9), 2275–2288.
Esterházy, D. et al., 2008. Production of reactive oxygen species by complex I (NADH: Ubiquinone Oxidoreductase) from Escherichia coli and comparison to the enzyme from mitochondria. Biochemistry 47 (12), 3964–3971.
Brosh, R.M., Jr., *Special Issue: DNA Helicases: Mechanisms, Biological Pathways, and Disease Relevance.* Genes, 2021. **12**(3): p. 356.
Mundra, S., Kjønaas, O.J., Morgado, L.N., Krabberød, A.K., Ransedokken, Y., Kauserud, H.J.F., 2021. Soil depth matters: shift in composition and inter-kingdom co-occurrence patterns of microorganisms in forest soils. FEMS Microbiol. Ecol. 97 (3).