7-7-2023

Comparative study of multiple approaches for identifying cultivable microalgae population diversity from freshwater samples

Amal A. Badr

Walid M. Fouad

RESEARCH ARTICLE

Comparative study of multiple approaches for identifying cultivable microalgae population diversity from freshwater samples

Amal A. Badr[1], Walid M. Fouad[1,2]*

1 Biotechnology Graduate Program, School of Sciences and Engineering, The American University in Cairo, New Cairo, Egypt, 2 Department of Biology, School of Sciences and Engineering, The American University in Cairo, New Cairo, Egypt

* wfouad@aucegypt.edu

Abstract

The vast diversity of microalgae imposes the challenge of identifying them through the most common and economical identification method, morphological identification, or through using the more recent molecular-level identification tools. Here we report an approach combining enrichment and metagenomic molecular techniques to enhance microalgae identification and identify microalgae diversity from environmental water samples. From this perspective, we aimed to identify the most suitable culturing media and molecular approach (using different primer sets and reference databases) for detecting microalgae diversity. Using this approach, we have analyzed three water samples collected from the River Nile on several enrichment media. A total of 37 microalgae were identified morphologically to the genus level. While sequencing the three-primer sets (16S rRNA V1-V3 and V4-V5 and 18S rRNA V4 region) and aligning them to three reference databases (GG, SILVA, and PR^2), a total of 87 microalgae were identified to the genus level. The highest eukaryotic microalgae diversity was identified using the 18S rRNA V4 region and alignment to the SILVA database (43 genera). The two 16S rRNA regions sequenced added to the eukaryotic microalgae identification, 26 eukaryotic microalgae. Cyanobacteria were identified through the two sequenced 16S rRNA regions. Alignment to the SILVA database served to identify 14 cyanobacteria to the genera level, followed by Greengenes, 11 cyanobacteria genera. Our multiple-media, primer, and reference database approach revealed a high microalgae diversity that would have been overlooked if a single approach had been used over the other.

Introduction

The advancement of next-generation sequencing (NGS) methods has made significant progress in microbiome research. Nevertheless, the gap in identifying the unknown content and "unculturable" communities and the necessity of the presence of live biological matter (the cultured microorganism) as a proof-of-concept test has allowed the culture methods to resurface. The utilization of the traditional microscopy techniques for directly studying of the environmental samples is subjective with the possibility of either overestimation or underestimation

of diversity if used as a sole identification method. Traditional microscopy techniques in microalgae identification depend on the level of experience of the person conducting it where small-sized cells are prone to either be completely missed or underrepresented [1]. Culturing samples before metagenomics has been utilized in various studies investigating the microbiota in different ecosystems; this approach is widely known as culturomics. Culturomics has been used for microbiome analysis of humans [2–8], marine sediments [9, 10], soil [11], root samples [12], and mouse gut [13]. Hence, we decided to utilize this approach in identifying microalgae diversity from freshwater environmental samples.

One of the approaches to identifying microalgae from environmental samples is the traditional morphological identification method, as mentioned above, through enrichment of the environmental sample before microscopy identification. The commonly used media for microalgae culture in the literature includes Blue-Green media (BG-11) and Bold's Basal media (BBM), which have been widely used to identify and maintain species diversity as they can sustain a wide range of species [14].

Microalgae identification can also be made through the molecular level analysis of conserved regions of the DNA, such as Ribulose Bisphosphate Carboxylase Large subunit gene (rbcL), 16S and 18S ribosomal RNA (rRNA) genes used for the profiling of prokaryotic and eukaryotic microalgae [15–17]. The reference database to align the sequenced results is equally important to the DNA barcodes sequenced. Hence, it is crucial to identify the most suitable database for microalgae identification. The SILVA database is predominantly used and is based on the small subunit (SSU) rRNAs that include the 16S and 18S rRNA in prokaryotes and eukaryotes, respectively. Therefore, they cover the three domains of life: Bacteria, Archaea and Eukarya. The SILVA database can assign taxonomic ranks to the genus level [18]. Greengenes (GG) database constitutes bacterial/archaeal 16S rRNA gene. The GG database is constructed from public aligned and checked databases for chimera [19, 20]. At the same time, the Protist Ribosomal Reference (PR^2) database contains sequences of the SSU rRNA and rDNA of the kingdom Eukaryota. The PR^2 database identifies eukaryotes through nuclear-encoded sequences, the 16S rRNA, present in the mitochondria or plastid of eukaryotic organisms [21]. Both GG and PR^2 provide taxonomic annotation down to the species level.

The influence of enrichment media on microalgae population growth and diversity in the literature remains elusive; however, still utilized [17, 22–24]. We previously investigated the efficiency of the commonly used BG-11 and BBM in microalgae enrichment for diversity identification and the effect of vitamin enrichment and reduced media nutrient on diversity enrichment. The results obtained using the reduced media provided insight into the limitation of both media and the need to investigate other media compositions [25]. Hence, this study aims to improve the detection and quantification of microalgae by comparing the usage of different enrichment media, DNA barcodes, and reference databases in the identification of cultivable microalgae diversity from an environmental sample. Our specific aims are: (1) increasing algal abundance above the detection threshold using four different enrichment media and (2) the utilization of multi-primer sets for SSU rRNAs (16S and 18S rRNA) and; (3) using different reference databases (SILVA, Greengenes (GG), and Protist Ribosomal Reference database (PR^2)) to evaluate their competence in the identification of microalgae.

Materials and methods

Study areas and sample enrichment

To investigate the competence of the different databases and the multi-primers studied, three water samples were collected from the Nile River in the Cairo governate (no specific permissions were required for sampling the three locations studied, which complied with all relevant

regulations.) (**S1 Table**) and enriched with different nutrient media. Water samples were collected 20 cm below the water surface from each location on the same day and kept in the fridge until transported and processed in the lab the next day. Portions of each sample were used for physicochemical analysis, and the remaining water samples were used to inoculate the nutrient media. Four different nutrient media were used for algal enrichment: Blue-Green medium (BG-11), Bold's Basal medium (BBM), half-strength Murashige and Skoog medium (MS), and a Modified medium (MM) that fits in the middle of the BG-11 and BBM in terms of the nitrogen and phosphorus composition (**S2 Table**). Each volume of nutrient media was inoculated with one-tenth of the environmental sample. For all tested factors, cultures were carried out in triplicate with continuous aeration and maintained for three weeks in a JSR-Growth Chamber 3-Side Illumination (model JSPC-960C2) at a temperature of 22±1°C. The light/dark regime was adjusted to 12h:12h, at a light intensity of 28 μmol photons m^{-2} s^{-1} [26]. Microalgae population growth was monitored by performing a cell count every four days using a Neubauer Improved Hemocytometer counting chamber for each replica under Zeiss AxioStar Plus light microscope at 400x magnification based on [27, 28] (**S1 Fig**).

Physicochemical analysis

The physicochemical analysis was performed to assess the composition of the microalgae's natural habitat identified in each location's water samples (**S3 Table**). The physical and chemical characteristics measured, including salinity, pH, and electrical conductivity (EC), were measured using a multimeter-probe (330i, WTW, Germany). The presence of cations and anions, such as sodium, chloride, sulfate, potassium, and calcium, were measured according to standard methods [29]. Total ammonia, nitrate, and phosphorus levels were evaluated according to a standard protocol [30].

Morphological identification of microalgae

Microalgae identification was based on the morphological characteristics observed under a bright-field microscope following the basic identification keys from references [31, 32]. The strains were examined under a light microscope (Leica) using the software LAS EZ (Leica DM500). Microalgae was identified at the genus level every seven days until the experiment was concluded.

DNA extraction and sequencing

Total genomic DNA (gDNA) was extracted using the PowerSoil DNA Isolation kit (MoBio, Carlsbad, CA) following the manufacturer's instruction with a minor alteration. Briefly, after 14 days of media enrichments, the cultured algae were collected from 200ml of each sample by centrifugation at 4000 rpm for 30 minutes. About 150–250 mg of the wet pellet was then snap-frozen by dropping the Eppendorf containing the wet pellet in liquid nitrogen. The frozen algae mass was ground using a pestle and used for gDNA extraction using the PowerSoil DNA Isolation kit following the manufacturer's instructions. The concentration and purity of the DNA extracted were confirmed using LVis Plate SPECTROstar® Nano (BMG LABTECH, UK), and integrity was confirmed using 1% agarose gel.

Three primer sets hyper-variable regions of 16S rRNA V1-V3 [33, 34], V4-V5 [33] and 18S rRNA V4 [35, 36] were used for sequencing (**S4 Table**). PCR and sequencing was performed at the Beijing Genomics Institute (BGI) (BGI-Tech, BGI-Shenzhen, Shenzhen, China). PCR amplification was performed using 2× Phanta Max Master kit (Vazyme Biotech Co., China) using 30ng gDNA as template. PCR conditions were as follows: initial denaturation at 94°C for 3 min; 30 cycles of 94°C for 30 s, annealing for 30s and the annealing temperature for each

primer set was 50˚C (16S V1-V3) and 53˚C (16S V4-V5 and 18S V4), extension at 72˚C for 45 s; and a final extension at 72˚C for 10 min. The PCR products were purified using AMPure XP beads, and the purified product was used to construct the DNA library. The prepared DNA library was used for sequencing. The sequencing was performed at the Beijing Genomics Institute (BGI) (BGI-Tech, BGI-Shenzhen, Shenzhen, China) using Illumina HiSeq 2500 platform to generate 250/300bp paired-end (PE) reads.

### Raw sequence processing and data analysis

Trimmed fastq files were filtered using MOTHUR v1.42.6 following the pipeline from the MiSeq standard operating procedure (SOP) available on their website (www.mothur.org/wiki/ MiSeq_SOP) [37]. Briefly, the fastq files were converted to fasta, and contigs were assembled using make.contigs. The files were screened using the MOTHUR command screen.seq. Reads selected were 457, 378, and 305bp fromthe 16S rDNA, V1-V3, V4-V5, and 18S rDNA sequences, respectively. Sequences with homopolymers longer than 8bp were discarded. The remaining sequences were de-replicated (unique.seqs) to merge duplicates and reduce the number of sequences to analyze. The remaining sequences were aligned to a database. The sequences of the 16S rDNA were aligned to the Greengenes database (Gg_13_8_99) and Protist Ribosomal Reference database (PR^2) [21] 18S rDNA was aligned to the SILVA database (Silva.nr_v132). The lowest level of the taxonomic hierarchy present in SILVA is the genus, while in Greengenes lowest taxonomic level that can be identified is the species level [18, 38]. Chimeras in the sequences were also filtered and removed (chimera.uchime and remove.seqs). Sequences were split into groups based on their taxonomy at the order level, and operational taxonomic units (OTUs) were assigned using the dist.seqs command. Alpha diversity were calculated based on normalized OTU abundance information obtained using the sample with the fewest sequences as a standard. The diversity indices calculated include Good's coverage, species observed (S$_{obs}$), Shannon's diversity index, Simpson's diversity index (InvSimpson), and Berger-Parker, and they were calculated using the MOTHUR SOP [39]. Non-metric multidimensional scaling (NMDS) was conducted with the Bray-Curtis index distance method for beta diversity index.The molecular data were further analyzed and visualized with R 3.6.1 packages ("vegan", "phyloseq", "taxa", and "ggplot2") executed via RStudio ("Open source and enterprise-ready professional software for data science-RStudio"; "R: The R Project for Statistical Computing").

Results

Overview of the freshwater microalgae taxa identified across different nutrient media

Pursuing the possibility of improving the media to better enrich the microalgae population and provide an evenness to the present diversity and enable the enrichment of rare species in the community. The growth rate of the algal population on the four different media tested indicated a comparable growth rate throughout the study. None of the media hindered the population growth; hence all were suitable for microalgae enrichment (S1 Fig).

The morphological assessment of the samples using the four different nutrient media aided in identifying 37 microalgae to the genus level. The highest relative abundance belonged to *Chlorella*, *Scenedesmus*, *Ankistrodesmus*, and *Selenastrum*. The highest diversity was identified when the sample was cultured on MM (36 genera), followed by BG-11 (31 genera), BBM (26 genera), and MS media (20 genera) (Fig 1A, S5 Table). The half-strength MS media had the lowest diversity, and all the genera identified through the MS media were identified with the
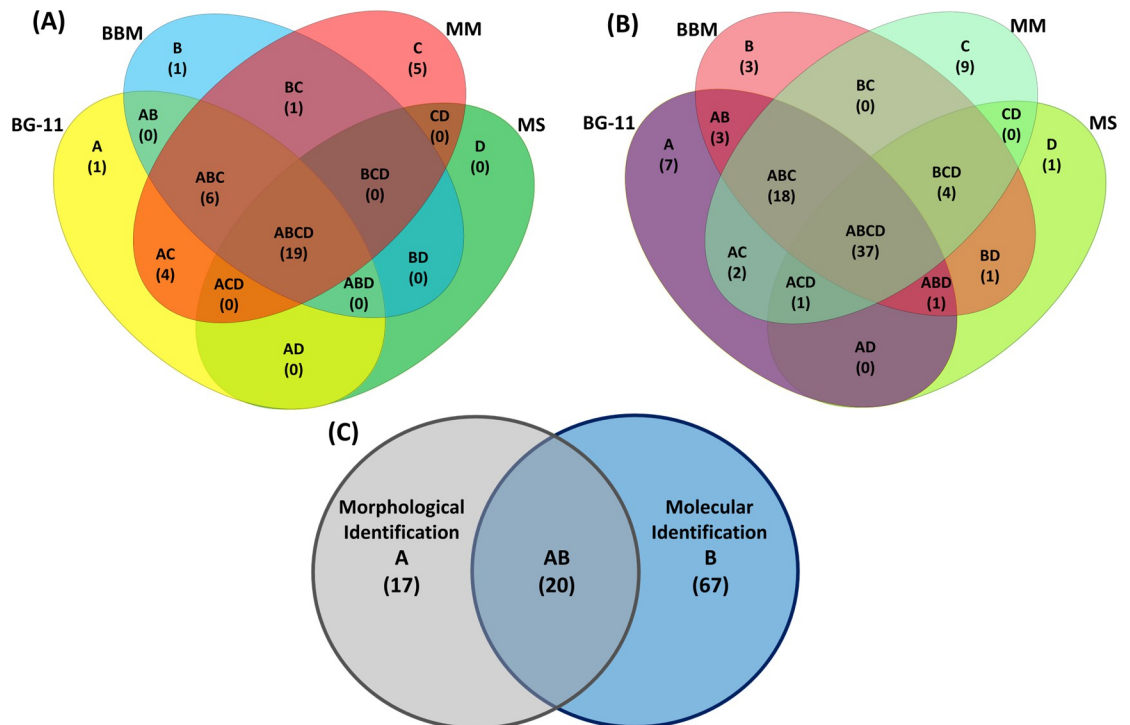
Fig 1. Freshwater microalgae taxa identified across different nutrient media. Venn diagram illustrating the number of shared and unique freshwater microalgae taxa across different culture media. The number of microalgae taxa identified on media BG-11, BBM, MM, and MS using two identification methods (A) Morphological identification, (B) Molecular-level identification using different DNA barcodes. (C) The total number of microalgae taxa commonly shared or uniquely identified using morphological and molecular identification methods. The genera identified are presented in S5 Table.

https://doi.org/10.1371/journal.pone.0285913.g001

other media. We decided to continue molecular analysis, including the MS medium, because there could be several limitations to morphological analysis and that includes, the volume sampled, the number of microscopic fields visualized from a single sample, and the possibility of human error. As the highest algal diversity was observed after 14 days, DNA extraction was conducted on all samples on day 14 [25]. Therefore, the DNA extraction for the molecular level analysis was conducted on the samples on the 14th day.

Using the three primer sets and sequencing through the Illumina HiSeq platform, the three databases identified 87 microalgae to the genus level. Similar to the results of the morphological analysis, the highest diversity was identified using the MM media (71 genera), followed by BG-11 (69 genera), BBM (67 genera), and MS media (45 genera) (Fig 1B). However, *Raphidonema* was only identified on half-strength MS media through molecular analysis.

By employing morphological and molecular methods, the sample enrichment media successfully aided in the identification of 104 microalgae at the genus level (Fig 1C and Table 1). The highest diversity was identified through the sequencing using the three primer sets sequencing sections of the hypervariable regions of 16S and 18S rRNA genes when compared to the morphological level identification. Seventeen genera were uniquely identified through the morphological level analysis, while the molecular-based approach identified 67 unique genera (Fig 1C).

Taxonomic composition and diversity comparison from different databases

The diversity of the microalgae cultured on the four media was analyzed through two regions of the 16S rRNA gene and one region in the 18S rRNA gene. The two 16S rRNA regions

Table 1. List of the algal genera identified in this study through both morphological and molecular- level.

Genus[Location]* (By Phyla)	Enrichment media*	Identification Method							
		Morphology	SILVA			GG		PR²	
			18S V4	16S V1-V3	16S V4-V5	16S V1-V3	16S V4-V5	16S V1-V3	16S V4-V5
Bacillariophyta									
Amphora [3]	A	-	-	-	-	-	-	+	-
Araphid pennate [1,2,3]	A B C D	-	-	-	-	-	-	-	+
Aulacoseira [1,2,3]	A B C	-	+	-	-	-	-	+	-
Cyclotella [1,2,3]	A B C D	+	+	-	-	-	-	+	-
Discostella [1,3]	A B C	-	+	-	-	-	-	-	-
Fistulifera [2]	A	-	-	-	-	-	-	+	-
Florenciella [2,3]	B C	-	-	-	-	-	-	+	-
Fragilaria [1,2]	A C D	-	-	-	-	-	-	+	-
Nannochloropsis [1,2,3]	A B C	-	-	-	-	-	-	-	+
Nitzschia [1,2,3]	A B C D	+	+	-	-	-	-	+	+
Ochromonas [1,2,3]	A B C D	-	+	-	-	-	-	-	+
Pedospumella [1,2,3]	A B C D	-	+	-	-	-	-	-	-
Polar Centric [1,2,3]	A B C D	-	-	-	-	-	-	+	+
Poterioochromonas [1,2]	B C D	-	+	-	-	-	-	-	-
Raphid pennate [1,2,3]	A B C D	-	-	-	-	-	-	+	+
Spumella [1,2,3]	A B C	-	+	-	-	-	-	-	-
Staurosira [2]	C	-	+	-	-	-	-	-	-
Stephanodiscus [2,3]	C	-	+	-	-	-	-	-	-
Thalassiosira [1,2,3]	A B C	-	+	-	-	-	-	+	-
Ulnaria [1,2,3]	A B C D	-	+	-	-	-	-	+	-
Charophyta/Streptophyta									
Chlorokybus [1,2]	A B C	-	-	-	-	-	-	+	+
Cosmarium [2]	A	+	-	-	-	-	-	+	-
Gonatozygon [2]	B	-	+	-	-	-	-	-	-
Mougeotia [1]	B	+	-	-	-	-	-	-	-
Staurastrum [1,2,3]	A B C	+	+	-	-	-	-	+	-
Staurodesmus [3]	B	-	+	-	-	-	-	-	-
Chlorophyta									
Acutodesmus [1,2,3]	A B C D	-	-	-	-	+	+	+	+
Anctinastrum [1,2,3]	A B C D	+	-	-	-	-	-	-	-
Ankistrodesmus [1,2,3]	A B C D	+	+	-	-	-	-	-	-
Asterococcus [1,2]	C	-	+	-	-	-	-	-	-
Botryococcus [2,3]	B	+	-	-	-	-	-	-	-
Carteria [1,2,3]	A B C D	-	+	-	-	-	-	-	-
Chlamydamonas [1,2,3]	A B C D	+	+	-	-	+	+	+	+
Chlorella [1,2,3]	A B C D	+	+	-	-	-	-	+	+
Chlorochytrium [2,3]	B C	-	+	-	-	-	-	-	-
Choricystis [1,2,3]	A B C D	-	+	-	-	-	-	-	-
Chlorococcum [3]	C	+	-	-	-	-	-	-	-
Closteriopsis [1,2,3]	A B C D	-	-	-	-	-	-	+	+
Coelastrum [1,2,3]	A B C D	+	+	-	-	-	-	-	-
Crucigenia [1,2,3]	A B C D	+	-	-	-	-	-	-	-
Desmodesmus [1,2,3]	A B C D	-	+	-	-	-	-	-	-
Dictyosphaerium [1,2,3]	A B C D	+	-	-	-	-	-	-	-

(Continued)

Table 1. (Continued)

Genus[Location]* (By Phyla)	Enrichment media*	Identification Method							
		Morphology	SILVA			GG		PR[2]	
			18S V4	16S V1-V3	16S V4-V5	16S V1-V3	16S V4-V5	16S V1-V3	16S V4-V5
Dunaliella [1,2,3]	A B C	-	-	-	-	+	-	+	-
Golenkinia [1,2,3]	A C	+	+	-	-	-	-	-	-
Gonium [1,2,3]	A B C D	+	-	-	-	-	-	+	+
Hydrodictyon [1,2,3]	A B C D	-	+	-	-	+	+	+	+
Koliella [1,2,3]	A B C D	-	-	-	-	-	-	-	+
Lagerheimia [1,2,3]	A B C D	+	-	-	-	-	-	-	-
Micractinium [1,2,3]	A B C D	+	+	-	-	-	-	-	-
Microglena [1,2,3]	A B C D	-	+	-	-	-	-	-	-
Monoraphidium [1,2,3]	A B C D	-	+	-	-	-	-	-	-
Mychonastes [1,2,3]	A B C	-	+	-	-	-	-	-	-
Neochlorosarcina [1,2,3]	B C D	-	+	-	-	-	-	-	-
Neodesmus [1,2]	A B C D	-	+	-	-	-	-	-	-
Oedogonium [1,2]	A B	-	-	-	-	-	-	+	+
Oocystis [1,2,3]	A B C D	+	+	-	-	-	-	+	+
Oophila [1,2,3]	A B C D	-	-	-	-	-	+	-	+
Parachlorella [1,2,3]	A B C D	-	-	-	-	-	-	+	-
Pectodictyon [2]	A	-	+	-	-	-	-	-	-
Pediastrum [1,3]	A B C	+	-	-	-	-	-	-	-
Pedinomonas [1,2,3]	A B C	-	+	-	-	-	-	-	-
Picochlorum [1,2,3]	A B C D	-	+	-	-	-	-	-	-
Pleodorina [1,2,3]	B C D	-	-	-	-	-	-	+	-
Raphidonema [1]	D	-	-	-	-	-	-	-	+
Scenedesmus [1,2,3]	A B C D	+	+	-	-	-	-	-	-
Selenastrum [1,2,3]	A B C D	+	-	-	-	-	-	-	-
Sphaeroplea [1,2,3]	A B C D	-	+	-	-	-	-	-	-
Tetracystis [1,2,3]	A B D	-	+	-	-	-	-	-	-
Tetradesmus [1,2,3]	A B C D	-	+	-	-	-	-	-	-
Tetraedron [1,2,3]	A B C D	+	-	-	-	-	-	-	-
Tetranephris [2]	C	-	+	-	-	-	-	-	-
Tetraselmis [1,3]	A B D	+	+	-	-	-	-	-	-
Tetraspora [1,2,3]	A B C D	+	-	-	-	-	-	-	-
Volvox [1,3]	A B	-	-	-	-	-	-	+	-
Cryptophyta									
Chroomonas [1,2,3]	A B	-	-	-	-	-	-	+	-
Cryptomonas [1,2,3]	A B C	+	+	-	-	-	-	+	+
Cyanobacteria									
Alkalinema [2]	B C D	-	-	+	+	-	-	-	-
Anabaena [1,3]	C D	+	-	-	-	-	+	-	-
Aphanizomenon [1,2,3]	A B C D	-	-	-	-	+	+	-	-
Chroococcus [1,2,3]	A B C D	+	-	-	-	-	-	-	-
Chrysosporum [1,2,3]	A B C	-	-	+	+	-	-	-	-
Coelosphaerium [1,2,3]	A C	+	-	-	-	-	-	-	-
Cyanobium [1,2,3]	A B C	-	-	+	+	-	-	-	-
Cylindrospermopsis [1,2,3]	A B C	-	-	+	+	-	-	-	-
Cylindrospermum [1,2,3]	A B C	+	-	-	-	+	+	-	-

(Continued)

Table 1. (Continued)

| Genus[Location]* (By Phyla)	Enrichment media*	Morphology	SILVA			GG		PR²	
			18S V4	16S V1-V3	16S V4-V5	16S V1-V3	16S V4-V5	16S V1-V3	16S V4-V5
Gloeothece [1]	*A*	-	-	-	-	-	-	+	-
Leptolyngbya [1,2,3]	*A B C D*	+	-	-	+	-	+	-	-
Limnothrix [1,2,3]	*A B C D*	-	-	+	+	-	-	-	-
Merismopedia [1,2,3]	*A B C D*	+	-	+	+	+	+	-	-
Mirocystis [1,2,3]	*A B C*	+	-	+	+	+	+	+	+
Nodosilinea [3]	*C*	-	-	-	+	-	-	-	-
Nostoc [1,2,3]	*A B C*	-	-	-	-	-	-	-	+
Phormidium [1]	*C*	-	-	+	-	-	-	-	-
Planktothricoides [1,2]	*A C*	-	-	+	+	+	+	-	-
Planktothrix [2]	*C*	-	-	-	-	+	-	-	-
Prochlorothrix [1,2]	*A C*	-	-	+	+	+	+	-	-
Psuedanabaena [1,2,3]	*A B C D*	+	-	+	+	+	+	-	-
Snowella [1,2,3]	*A B C*	-	-	+	+	+	+	-	-
Spirulina [1,2,3]	*A B C D*	+	-	-	-	-	-	-	-
Synechococcus [1,2,3]	*A B C*	-	-	-	-	+	+	+	+
Dinoflagellata									
Heterocapsa [1,2,3]	*A B C D*	-	-	-	-	-	-	-	+
Thoracosphaera [1,2,3]	*A B C D*	-	-	-	-	-	-	+	+
Euglenophyta									
Euglena [1,2,3]	*A B C D*	+	-	-	-	-	-	+	+
Ochrophyta									
Botrydiopsis [1,2,3]	*A B C D*	+	-	-	-	-	-	-	-

*Genera that were identified through both morphological and molecular level through the four-enrichment media tested: A = BG-11, B = BBM, C = MM, and D = MS. Locations are annotated by number superscribed beside genera name: [1]Nile1, [2]Nile2 and [3]Nile3.

sequenced (V1-V3 and V4-V5) combined had a total of 852,517 sequences. In comparison, the 18S rRNA V4 region had 423,623 sequences. After the filtering and processing, the two 16S rRNA regions combined had 271,511 unique sequences, and the 18S rRNA region had 344,887 sequences. The 16S rRNA OTUs were assigned by alignment to three reference databases: SILVA, Greengenes (GG), and the Protist Ribosomal Reference (PR²). The 18S rRNA sequences were aligned to the SILVA database.

5298 taxa were identified using the 16S V1-V3 and SILVA as a reference database, while using the GG and the PR² database, 5498 and 5942 taxa were classified, respectively. Moreover, 397 taxa were classified using the 18S rRNA through the SILVA database. Further filtering was performed to eliminate unclassified phyla to focus on the classified microalgae phyla/genera using RStudio packages.

Cyanobacteria identification. The 16S rRNA hypervariable regions V1-V3, and V4-V5 and three databases identified 18 and 17 genera, respectively (**Fig 2A and 2B**). By analyzing the two SSU 16S rRNA gene regions sequenced, the SILVA database aided in identifying higher diversity of cyanobacteria (14 genera) over the other tested databases (**Fig 3A and 3D**). Alignment to SILVA alone, both 16S rRNA regions produced identical results, and either region would be sufficient for cyanobacteria identification. Certain cyanobacteria were only identified through GG and PR² databases (**Fig 3B and 3E**). Using the GG database, two genera of cyanobacteria, *Aphanizomenon* and *Planktothrix*, were identified exclusively through the
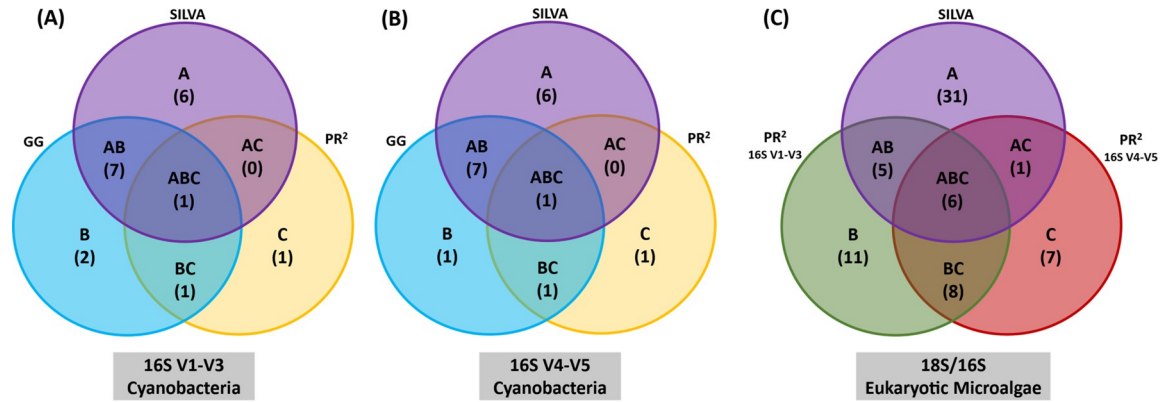
Fig 2. Venn diagrams illustrating the unique and shared microalgae identified through the different databases (SILVA, GG, and PR²) and the different DNA barcodes sequenced. Cyanobacteria genera were identified through (A) 16S V1-V3 region and (B) 16S V4-V5 rRNA region. (C) Eukaryotic genera identified through 18S rRNA gene annotated using SILVA database and 16S rRNA gene annotated using PR² database.

V1-V3 and V4-V5 regions of the 16S rRNA gene and were not identified using SILVA and PR² (**Fig 3**). Not much was expected from the PR² database and identifying cyanobacteria (**Fig 3C and 3F**); however, two cyanobacteria genera were only identified through the PR² database. Using the V1-V3 region, *Gloeothece* was identified in the BG-11 medium, and the V4-V5 region revealed *Nostoc* in the nutrient media BG-11, BBM, and MM. Nonetheless, using the primers V1-V3 and V4-V5 helped to uncover more cyanobacteria using the SILVA and GG databases.
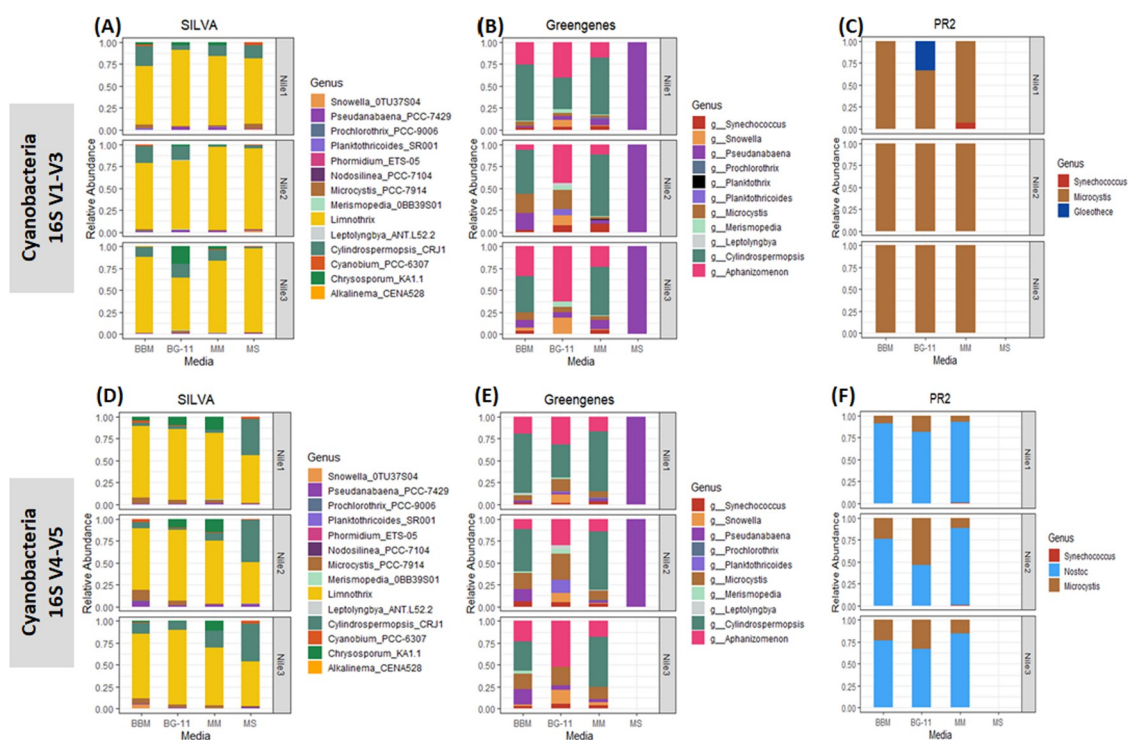
Fig 3. Cyanobacteria identification. Relative abundance of cyanobacteria identified through 16S rDNA V1-V3 (top-half) and 16S rDNA V4-V5 region (bottom-half) annotated through (A & D) SILVA, (B & C) GG and (C & F) PR² databases.
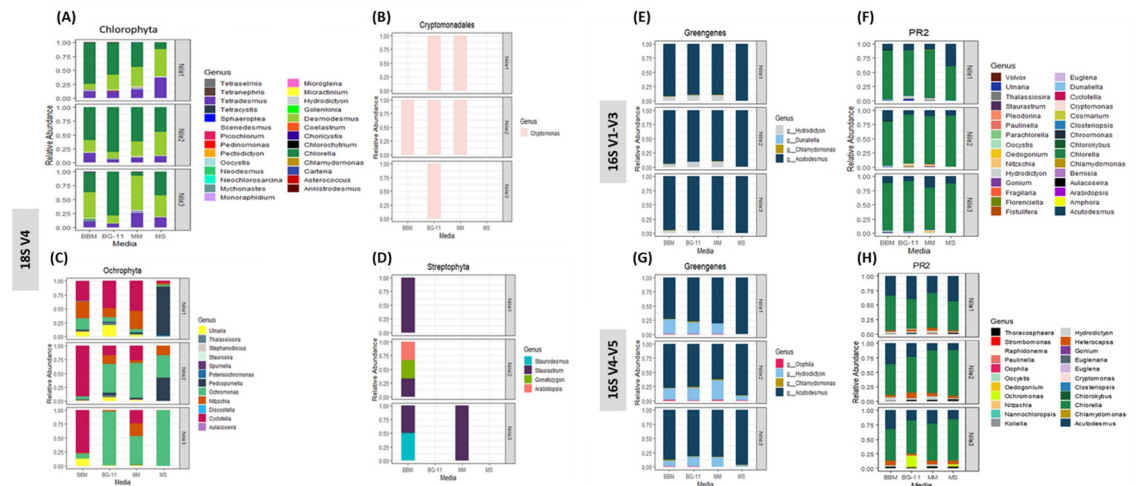
Eukaryotic microalgae identification. Both the 16S and the 18S rRNA sequenced regions identified 68 eukaryotic microalgae (Fig 2C). The SILVA database and the 18S rRNA V4 region revealed the highest diversity of 43 eukaryotic microalgae, with 27 genera belonging to the phylum Chlorophyta and 12 genera belonging to the Ochrophyta. The remaining genera belong to the phylum Streptophyta (3 genera) and Cryptophyta (1 genus) (Fig 4A–4D). The PR^2 database and the two 16S rRNA genes classified 37 eukaryotic microalgae genera. The 16S V1-V3 fragment divulged 28 genera, out of which 10 genera were uniquely identified by the PR^2 database. In contrast, the 16S V4-V5 fragment alignment identified 22 microalgae genera (Fig 4E–4H). Using only the SILVA database and the 18S rRNA V4, 25 eukaryotic microalgae would have been missed from the study as these genera were only revealed through the sequencing of the 16S V1-V3 and the V4-V5 regions and annotated using the PR^2.

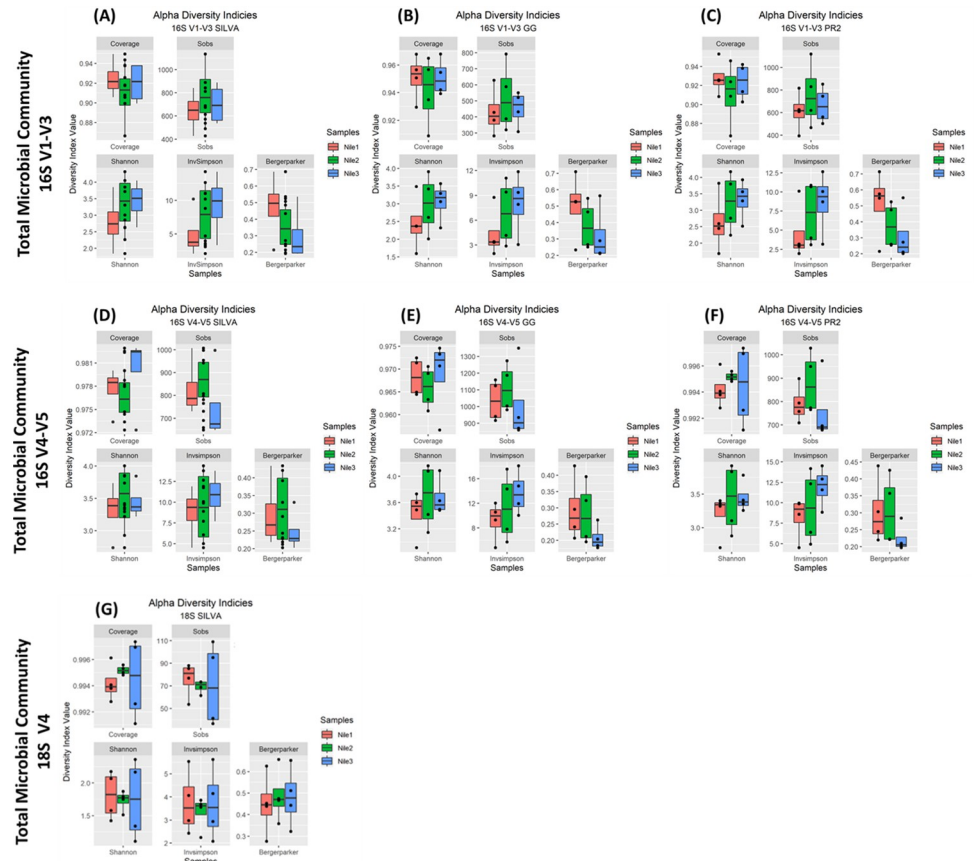## Alpha diversity indices of total microbial community based on OTUs

The enrichment of the microalgae diversity through nutrient media has created a closed eco-system, and hence the alpha diversity indices are applied to the new community. The diversity indices here are used to validate the enrichment of the microalgae community and do not reflect their natural relative abundance. The 16S V1-V3 region through the three databases had similar Good's coverage index ranging from 94–95% across the samples using the GG database, while the PR^2 coverage index ranged from 91–93% (Fig 5A–5C). The observed richness (S_{obs}) in terms of OTUs was the highest in the Nile2 sample, followed by the Nile3 and Nile1 samples across the three databases (Fig 5A–5C). The V4-V5 region of the 16S rDNA had a higher Good's coverage (<96%) when calculated across the three databases. The S_{obs} are high when analyzed through the V4-V5 region over the V1-V3 region. However, the Nile2 still had the highest S_{obs}, followed by Nile1 and Nile3 using both 16S regions (Fig 5D–5F). The Shannon index for Nile2 was the highest amongst the other Nile samples when analyzing the sequences for the V4-V5 region (Fig 5D–5F), while the higher Shannon index was observed in Nile3 when examining the V1-V3 region (Fig 5A–5C), followed by Nile2 and Nile1 samples.

The high InvSimpson index reflects the increase in diversity. The InvSimpson values increase is observed across Nile3 samples across all the sequenced regions, followed by Nile2

Fig 4. Eukaryotic microalgae identification. Relative abundance of eukaryotic microalgae identified through the different databases (SILVA, GG, and PR^2) and the different DNA barcodes sequenced. (A-D) are eukaryotic microalgae identified through 18S rRNA gene annotated using SILVA database. Relative abundance of eukaryotic microalgae genera identified using 16S V1-V3 through Greengenes and PR^2 (E & F), respectively, 16S V4-V5 rRNA gene (G & H) Greengenes and PR^2, respectively.

https://doi.org/10.1371/journal.pone.0285913.g004

Fig 5. Alpha diversity indices of total microbial community based on OTUs. The total calculated alpha diversity are as follows: Good's Coverage; S_{obs} (total number of OTUs observed); Shannon; InvSimpson and Breger-Parker. The first half belonging to the V1-V3 region (A-C) and the second-half represents the V4-V5 region (D-F), and last plot belongs to 18S V4. (A, D) SILVA, (B, E) GG, and (C, F) PR^2 database. (G) Alpha diversity indices calculated through 18S rRNA gene and SILVA.

https://doi.org/10.1371/journal.pone.0285913.g005

and Nile1 samples. However, the 16S V4-V5 and the 18S V4 of Nile1 and Nile2 have a similar median (**Fig 5D–5G**). The final alpha diversity index evaluated regarding the OTUs abundance is the Berger-Parker index. The decrease in this index indicates an increase in diversity. The Nile3 sample has the lowest Berger-Parker index indicating the presence of high diversity followed by Nile2 and Nile3 across all the databases and regions (**Fig 5**). This finding is in line with the other two calculated indices, Shannon and InvSimpson, for the three locations. The 18S V4 region for eukaryotic microorganisms analyzed through the SILVA database had a Good's coverage >99%, while the highest S_{obs} was observed in the Nile1 sample followed by Nile3 and Nile2. However, Shannon and InvSimpson indices across all the samples were similar Nile1 (1.81, 3.75), Nile2 (1.73, 3.33), and (1.74, 3.69) (Shannon, InvSimpson, respectively) (**Fig 5G**). The Berger-Parker index was also similar for all three samples 0.45, 0.49, and 0.48 for Nile1, Nile2, and Nile3. The diversity indices indicate that the Nile3 is the most abundant and with an even distribution of genera and the absence of the dominance of a genus. This is followed by the Nile2 and Nile1 samples. Regarding the 18S V4 region annotated by SILVA, the diversity indices indicate that Nile1 has the richest diversity with a high evenness between the genera identified. This is followed by Nile3 and then Nile2.
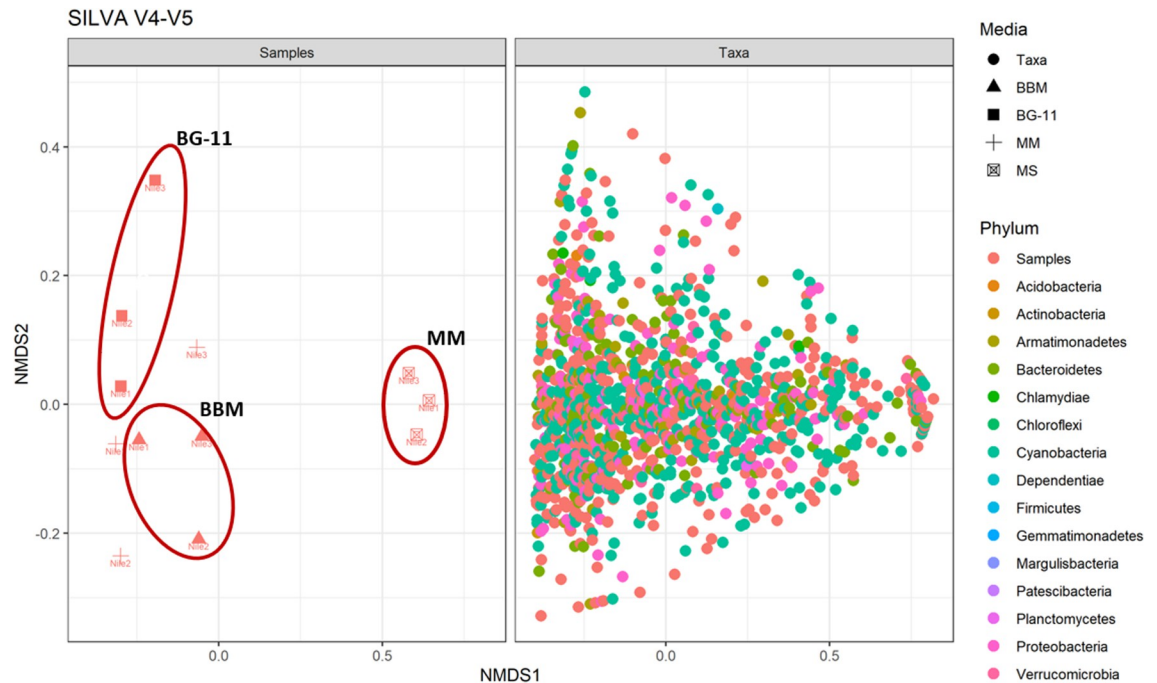
Fig 6. Non-metric multidimensional scaling analysis (nMDS) computed on Bray-Curtis similarity index obtained for microbial community-identified using SILVA 18S V4, water samples, and media as factors.

https://doi.org/10.1371/journal.pone.0285913.g006

The relationship between identified phyla, media, and location was examined through the Bray-Curtis similarity index based on the microbial composition identified through the results of the SILVA 18S V4 as it identified the highest overall diversity. An unexpected clustering pattern was observed, and samples were clustered with samples enriched on the same media rather than being clustered by their corresponding locations (**Fig 6**). This result demonstrates that using only one medium can lead to the enrichment of certain microalgae species while missing others, resulting in misleading data concerning the diversity of studied environmental samples.

Discussion

This study aimed to evaluate culture-based morphology and the multi-primer/multi-database molecular approach in identifying microalgae diversity from an environmental sample, in our case, the River Nile. The challenge presented was cultivating the highest possible microalgae diversity and identifying them microscopically and through different DNA barcodes. Each of the four nutrient media tested had all the necessary macronutrients (N, P, K, and S) and micronutrients (such as Fe, Cu, Mn, Zn, Co, and Mo) to support microalgae growth. However, each nutrient media contained different macro and micronutrient concentrations. Differences in nutrients and media components could have led to the preference of certain microalgae growth over others, such as nitrogen, which differs among the different media utilized in form and concentration [40–42]. BG-11, BBM, and MM media contain sodium nitrate as the main nitrogen source. The MS medium contains the highest nitrogen concentration in the form of ammonium and potassium nitrate. Microalgae use nitrogen in different forms, such as nitrate, nitrite, ammonium, as inorganic nitrogen sources, and urea, as an organic nitrogen source. The different nitrogen sources are first reduced to ammonium which is converted/integrated into amino acids through different pathways in microalgae [42]. While ammonium is a

preferred nitrogen source due to its low-energy assimilation, sodium, and potassium nitrates are still preferred in media over ammonium. This is attributed to ammonium ions (NH_4^+) being converted to ammonia gas (NH_3) under aeration at alkaline conditions and thus considered a loss of nitrogen source [43]. Hence, this may be a reason why MS media had the lowest identified diversity even though it was rich in nitrogen content.

Citric acid and EDTA are other critical media components. The citric acid present in BG-11 and MM media is responsible for solubilizing salt components in the media and preventing iron precipitation, making it readily available to the microalgae, thus resulting in an enhanced growth rate. Iron is considered an essential limiting micronutrient [43]. Therefore, it can be a leading cause of the higher diversity in MM and BG-11 media. BG-11, BBM and MM media also have EDTA, disodium salt acting as a chelating agent. In contrast, MS medium only has its iron chelated in the form of Ferric sodium EDTA, which has been reported to be a better source of iron [44, 45]. Yet, the MS medium did not uniquely enrich any microalgae over the other tested media when morphologically analyzed. This indicated that the ratio of iron is relatively low compared to the other media, and its increase could have resulted in a higher diversity enrichment.

pH can also increase due to the microalgae uptake of inorganic carbon, and photosynthesis increases O_2 production, which in turn may increase the pH of the culture [46, 47]. In our experiments nutrient content consumption was not measured. Nutrient depletion and accumulation of waste products from the enriched microalgae biomass is assumed after 14 days and can lead to the increase in pH. After 14 days the intermediate and rare microalgae (low in abundance) are further reduced, boosting the dominant and tolerant microalgae such as *Ankistrodesmus*, *Chlorella*, and *Scenedesmus* [48, 49]. This is common as green algae are known to dominate and reduce the growth of cyanobacteria [50].

To perform a molecular-level analysis at time-point zero, it would have been necessary to filter the samples to gather the existing microbiome. However, this filtration process could have caused the loss of picoplankton [51] or dormant cysts, potentially resulting in the omission of certain species [52]. We recognize that conducting molecular analysis on the three samples before enrichment on day zero would have further enhanced our understanding of the algal population dynamics from the beginning. However, filtration of the samples could have resulted in the loss of picoplankton *Choricystis*, *Nannochloropsis*, *Picochlorum*, *Synechococcus*, and *Thalassiosira*. The identification of diversity is significantly influenced by filtration procedures, which involve the selection of filter paper type, pore size, and the amount of environmental sample filtered [53–55]. Given the potential limitations associated with filtration, we decided to proceed with enrichment.

Sixty-seven of 104 microalgae genera were only identified through sequencing and molecular-level identification while only 17 unique genera were identified morphologically. It is possible that the discrepancy is due to the fact that the species had undergone cell lysis either in the environment or during the culturing period during the experiment releasing their genetic material [1] or were in a dormant phase [52]. Accordingly, the DNA of the species has been detected through sequencing and not identified morphologically [56–58]. Moreover, some species may have been missed through microscopy as the method depends on the scientist and their skill in identifying microalgae through morphology [59, 60].

The 17 unique genera identified morphologically were not identified through sequencing. This can be explained by the possibility that some microalgae were incorrectly classified through morphological characteristics due to the close resemblance to other microalgae in the same phylum. For example, *Tetraspora* (Chlorophyta) can be confused with *Chlamydomonas* or *Gloeococcus* while other genera such as *Leptolyngbya* and *Spirulina* (Cyanobacteria) can be incorrectly classified with other filamentous Cyanobacteria [61]. The presence of high

nutrients and asexual reproduction can hinder the morphological identification of unicellular microalgae due to their phenotypic plasticity under nutrient and environmental alteration [52, 62].

The number of unique microalgae identified through the molecular-level identification can be higher due to the presence of synonyms of genera between databases that may cause a double count between morphological and molecular-level identification or result in a "hit and miss" between databases [63]. *Aphanizomenon* identified through 16S rRNA primers using GG is considered as a basionym of *Chrysosporum* which was only identified through SILVA using the 16S rRNA primers [64]. *Chrysosporum* has a homotypic synonym *Anabaena bergii* [65, 66] which was identified through microscopy to the genus level as *Anabaena* using MM and MS media in Nile1 but using the two 16S rRNA primer sets and alignment against SILVA database. Other basionym include *Raphidonema* and *Koliella* (Taxonomy ID: 153998 and 33092) [67].

A drawback of the databases used in this study is that they are not specific to microalgae and are not constantly updated with microalgal sequences compared to microbial taxonomies such as in studies on the gut and oral microbiomes or general microbial communities in environmental samples [38, 68–72].

Genera such as *Selenastrum* identified through morphological-level identification has a synonym *Monophidium* profiled through SILVA 18S V4 region. Even though there is a complete sequence entry of the 18S rRNA gene of *Selenastrum* present in the SILVA database along with a sequence for *Monophidium* [73]. This limitation can be ascribed to the SILVA database being a curated database, and there are several entries with different synonyms of the same genera [74]. However, some of the microalgae identified only on a morphological level have characteristic features that can be easily distinguished microscopically and are not mistaken with other genera such as *Actinastrum*, *Coelastrum*, *Dictyosphaerium*, *Largerheimia*, *Tetraedron*, *Pediastrum*, and *Selenastrum* [31, 32]. Nevertheless, these species were not identified on a molecular level through any of the regions analyzed or through the three reference databases used. The complete 18S rRNA gene of the genera *Actinastrum*, *Coelastrum*, *Dictyosphaerium*, *Tetraedron*, *Pediastrum*, and *Selenastrum* are present in the SILVA database, while *Largerheimia* 18S rRNA gene was not available on SILVA but is present in a different database ENA (European Nucleotide Archive) that has not been used in this study for the analysis of the 18S rRNA sequences [73].

BG-11, BBM and MM were the most suitable for culturing and enrichment of the highest diversity from an environmental sample, compared to MS media that did not identify any unique species. The enrichment using culture-based methods can be considered an alternative to the more expensive approach of increasing the number of sequence reads to detect the less abundant species due to limited detection thresholds. The culture-based method followed by the molecular level using several different hypervariable regions eliminates these biases resulting from using one medium or a sequencing region and broadens the identification level. It is critical to recognize that different enrichment media and sequencing regions complements one another and provides a more complete diversity profile.

The combined use of both the culture-based method and sequence analysis of multiple rRNA regions/genes (targeting both nuclear and plastid genomes) [75, 76] facilitates a more reliable and comprehensive approach to identifying total microalgae biodiversity. This approach can be considered promising as it targets microalgae identification from an environmental sample compared to the commonly used metagenomic sequencing that utilizes the environmental sample directly. These are usually high in microbial biomass and low in microalgae biomass. We acknowledge that the culturomics approach to identifying microalgae from environmental studies needs further improvement and further studies are needed to improve

the method and further define it for reproducibility. This approach to identifying microalgae can be further explored using samples from different ecosystems.

Conclusion

Four enrichment media were employed in this study to examine their efficacy in enhancing the diversity of microalgae obtained from the environment. The two 16S rDNA regions and one 18S rDNA region were then tested and aligned to various databases to determine the best alternative methods for identifying microalgae to classical microscopy. Overall, we advise using both morphological and morphological approaches since dependence on one technique over another can be restrictive and result in genera omission. When relying on molecular-level identification, the reference database and the DNA barcode employed must be considered from several prospective. The SILVA database was the most effective in identifying cyanobacteria and both 16S rDNA regions produced identical results. Whereas, Greengenes database identified less cyanobacteria genera compared to SILVA using the 16S rDNA Most of the detected eukaryotic microalgae could be recognized mainly using the 18S V4 region and alignment to SILVA, followed by the alignment of the 16S V1–V3 to the PR2. Therefore, we recommend coupling enrichment media (BG-11, BBM, and MM) with morphological level identification, followed by molecular level identification using the 16S V1-V3 or V4-V5 region and 18S V4 region for both cyanobacteria and eukaryotic microalgae identification.

Supporting information

S1 Fig. Growth curve across all tested media to ensure enrichment and media suitability uniform algal growth.
(TIF)

S1 Table. The geographic coordinate for environmental samples collected from the river Nile and used in this study.
(PDF)

S2 Table. The chemical composition (mg L^{-1}) for the four growth media used in the study.
(PDF)

S3 Table. Physiochemical analysis of collected samples.
(PDF)

S4 Table. Primer sequences used for PCR amplification and sequencing.
(PDF)

S5 Table. Unique and shared genera identified across the four different media A = BG-11, B = BBM, C = MM and D = MS. (a) morphologically across the four media, (b) molecularly using the three primer- sets and the three reference databases, and (c) combining both molecular and morphological approaches.
(PDF)

Acknowledgments

Author Contributions

Conceptualization: Amal A. Badr, Walid M. Fouad.

Data curation: Amal A. Badr.

Funding acquisition: Amal A. Badr.

Investigation: Amal A. Badr.

Methodology: Amal A. Badr, Walid M. Fouad.

Resources: Walid M. Fouad.

Supervision: Walid M. Fouad.

Visualization: Amal A. Badr.

Writing – original draft: Amal A. Badr.

Writing – review & editing: Walid M. Fouad.

References

1. Xiao X, Sogge H, Lagesen K, Tooming-Klunderud A, Jakobsen KS, Rohrlack T. Use of High Throughput Sequencing and Light Microscopy Show Contrasting Results in a Study of Phytoplankton Occurrence in a Freshwater Environment. PLoS One [Internet]. 2014 Aug 29 [cited 2022 Apr 28]; 9(8):e106510. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0106510 PMID: 25171164

2. Kaspar U, Kriegeskorte A, Schubert T, Peters G, Rudack C, Pieper DH, et al. The culturome of the human nose habitats reveals individual bacterial fingerprint patterns. Environ Microbiol [Internet]. 2016 Jul 1 [cited 2022 Apr 27]; 18(7):2130–42. Available from: https://pubmed.ncbi.nlm.nih.gov/25923378/ https://doi.org/10.1111/1462-2920.12891 PMID: 25923378

3. Whelan FJ, Waddell B, Syed SA, Shekarriz S, Rabin HR, Parkins MD, et al. Culture-enriched metagenomic sequencing enables in-depth profiling of the cystic fibrosis lung microbiota. Nature Microbiology 2020 5:2 [Internet]. 2020 Jan 20 [cited 2022 Apr 27];5(2):379–90. Available from: https://www.nature.com/articles/s41564-019-0643-y https://doi.org/10.1038/s41564-019-0643-y PMID: 31959969

4. Zou Y, Xue W, Luo G, Deng Z, Qin P, Guo R, et al. 1,520 reference genomes from cultivated human gut bacteria enable functional microbiome analyses. Nature Biotechnology 2019 37:2 [Internet]. 2019 Feb 4 [cited 2022 Apr 27];37(2):179–85. Available from: https://www.nature.com/articles/s41587-018-0008-8 https://doi.org/10.1038/s41587-018-0008-8 PMID: 30718868

5. Amrane S, Hocquart M, Afouda P, Kuete E, Pham TPT, Dione N, et al. Metagenomic and culturomic analysis of gut microbiota dysbiosis during Clostridium difficile infection. Sci Rep. 2019 Dec 1; 9(1). https://doi.org/10.1038/s41598-019-49189-8 PMID: 31488869

6. Naud S, Khelaifia S, Mbogning Fonkou MD, Dione N, Lagier JC, Raoult D. Proof of Concept of Culturomics Use of Time of Care. Front Cell Infect Microbiol. 2020 Nov 23; 10:714. https://doi.org/10.3389/fcimb.2020.524769 PMID: 33330116

7. Greub G. Culturomics: a new approach to study the human microbiome. Clinical Microbiology and Infection [Internet]. 2012 Dec 1 [cited 2022 May 4]; 18(12):1157–9. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/1469-0691.12032 PMID: 23148445

8. Lagier JC, Dubourg G, Million M, Cadoret F, Bilen M, Fenollar F, et al. Culturing the human microbiota and culturomics. Nature Reviews Microbiology 2018 16:9 [Internet]. 2018 Jun 24 [cited 2022 May 7]; 16 (9):540–50. Available from: https://www.nature.com/articles/s41579-018-0041-0 https://doi.org/10.1038/s41579-018-0041-0 PMID: 29937540

9. Mu DS, Liang QY, Wang XM, Lu DC, Shi MJ, Chen GJ, et al. Metatranscriptomic and comparative genomic insights into resuscitation mechanisms during enrichment culturing. Microbiome [Internet]. 2018 Dec 26 [cited 2022 Apr 27]; 6(1):1–15. Available from: https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-018-0613-2

10. Mu DS, Ouyang Y, Chen GJ, Du ZJ. Strategies for culturing active/dormant marine microbes. Mar Life Sci Technol. 2021 May 1; 3(2):121–31. https://doi.org/10.1007/s42995-020-00053-z PMID: 37073338

11. Costa OYA, Costa OYA, de Hollander M, Pijl A, Liu B, Kuramae EE. Cultivation-independent and cultivation-dependent metagenomes reveal genetic and enzymatic potential of microbial community involved in the degradation of a complex microbial polymer. Microbiome [Internet]. 2020 Jun 1 [cited 2022 Apr 27]; 8(1):1–19. Available from: https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-020-00836-7

12. Zhang J, Liu YX, Guo X, Qin Y, Garrido-Oter R, Schulze-Lefert P, et al. High-throughput cultivation and identification of bacteria from the plant root microbiota. Nature Protocols 2021 16:2 [Internet]. 2021 Jan 13 [cited 2022 Apr 27]; 16(2):988–1012. Available from: https://www.nature.com/articles/s41596-020-00444-7 https://doi.org/10.1038/s41596-020-00444-7 PMID: 33442053

13. Liu C, Zhou N, Du MX, Sun YT, Wang K, Wang YJ, et al. The Mouse Gut Microbial Biobank expands the coverage of cultured bacteria. Nature Communications 2020 11:1 [Internet]. 2020 Jan 7 [cited 2022 Apr 27]; 11(1):1–12. Available from: https://www.nature.com/articles/s41467-019-13836-5 https://doi.org/10.1038/s41467-019-13836-5 PMID: 31911589

14. Khaw YS, Khong NMH, Shaharuddin NA, Yusoff FM. A simple 18S rDNA approach for the identification of cultured eukaryotic microalgae with an emphasis on primers. J Microbiol Methods. 2020 May 1; 172:105890. https://doi.org/10.1016/j.mimet.2020.105890 PMID: 32179080

15. Hadi SIIA, Santana H, Brunale PPM, Gomes TG, Oliveira MD, Matthiensen A, et al. DNA Barcoding Green Microalgae Isolated from Neotropical Inland Waters. PLoS One [Internet]. 2016 Feb 1 [cited 2022 May 8]; 11(2):e0149284. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0149284 PMID: 26900844

16. Wang C, Zou S, Fei C, Wang C, Gao Z, Bao Y, et al. How DNA barcoding can be more effective in microalgae identification: a case of cryptic diversity revelation in Scenedesmus (Chlorophyceae). Scientific Reports 2016 6:1 [Internet]. 2016 Nov 9 [cited 2022 May 8]; 6(1):1–13. Available from: https://www.nature.com/articles/srep36822 https://doi.org/10.1038/srep36822 PMID: 27827440

17. Ballesteros I, Terán P, Guamán-Burneo C, González N, Cruz A, Castillejo P. DNA barcoding approach to characterize microalgae isolated from freshwater systems in Ecuador. Neotrop Biodivers [Internet]. 2021 [cited 2022 May 8]; 7(1):170–83. Available from: https://www.tandfonline.com/doi/abs/10.1080/23766808.2021.1920296

18. Balvočiute M, Huson DH. SILVA, RDP, Greengenes, NCBI and OTT—how do these taxonomies compare? BMC Genomics. 2017 Mar 14;18.

19. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a {Chimera}-{Checked} {16S} {rRNA} {Gene} {Database} and {Workbench} {Compatible} with {ARB}. Appl Environ Microbiol. 2006 Jul; 72(7):5069–72. https://doi.org/10.1128/AEM.03006-05 PMID: 16820507

20. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, et al. An improved {Greengenes} taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J. 2012 Mar; 6(3):610–8. https://doi.org/10.1038/ismej.2011.139 PMID: 22134646

21. Guillou L, Bachar D, Audic S, Bass D, Berney C, Bittner L, et al. The Protist Ribosomal Reference database (PR2): A catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. Nucleic Acids Res [Internet]. 2013 Jan 1 [cited 2021 Mar 20]; 41(D1):D597. Available from: /pmc/articles/PMC3531120/ https://doi.org/10.1093/nar/gks1160 PMID: 23193267

22. Kulichová J, Škaloud P, Neustupa J. Molecular diversity of green corticolous microalgae from two sub-{Mediterranean} {European} localities. Eur J Phycol. 2014 Jul; 49(3):345–55.

23. Sánchez Roque Y, Pérez-Luna Y, Moreira Acosta J, Farrera Vázquez N, Berrones Hernández R, Saldaña Trinidad S, et al. Evaluation of the population dynamics of microalgae isolated from the state of {Chiapas}, {Mexico} with respect to the nutritional quality of water. BDJ. 2018 Sep; 6:e28496.

24. Wu Y, Guan K, Wang Z, Xu B, Zhao F. Isolation, {Identification} and {Characterization} of an {Electrogenic} {Microalgae} {Strain}. Bansal V, editor. PLoS One. 2013 Sep; 8(9):e73442. https://doi.org/10.1371/journal.pone.0073442 PMID: 24019922

25. Badr AA, Fouad WM. Identification of culturable microalgae diversity in the River Nile in Egypt using enrichment media. African Journal of Biological Sciences [Internet]. 2021 Apr 5 [cited 2021 May 16]; 3 (2):50. Available from: https://www.afjbs.com/files/1618246323_(7)_AFJBS20210120_(p_50-64).pdf

26. Maltsev Y, Maltseva K, Kulikovskiy M, Maltseva S. Influence of Light Conditions on Microalgae Growth and Content of Lipids, Carotenoids, and Fatty Acid Composition. Biology (Basel) [Internet]. 2021 Oct 1 [cited 2022 May 3]; 10(10). Available from: /pmc/articles/PMC8533579/ https://doi.org/10.3390/biology10101060 PMID: 34681157

27. Karlson B, Godhe A, Cusack C, Bresnan E. Microscopic and molecular methods for quantitative phytoplankton analysis. In: Intergovernmental Oceanographic Commission Manuals and Guides 55. 2010. p. 5–12.

28. Raeesossadati MJ, Ahmadzadeh H, McHenry MP, Moheimani NR. CO2 bioremediation by microalgae in photobioreactors: Impacts of biomass and CO2 concentrations, light, and temperature. Algal Res [Internet]. 2014 Oct [cited 2019 Aug 7]; 6:78–85. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2211926414000903

29. ASTM Standard D4327. Test Method for Anions in Water by Suppressed Ion Chromatography. 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428–2959, United States; 2017.

30. Rice EW, Franson MAH. 21st Edition 2005. 21. ed., c. Eaton AD, Association APH, Association AWW, Federation WPC, editors. Washington, DC: American Public Health Assoc; 2005.

31. Bellinger EG, Sigee DC. Freshwater Algae: Identification and Use as Bioindicators [Internet]. Freshwater Algae: Identification and Use as Bioindicators. Chichester, West Sussex, UK; Hoboken, NJ: Wiley-Blackwell; 2010. Available from: http://public.eblib.com/choice/publicfullrecord.aspx?p=530045

32. Evans RI, Prescott GW. How to Know the Fresh-Water Algae. [Internet]. Bulletin of the Torrey Botanical Club. W.C. Brown Company; 1956. 311 p. (Pictured key nature series; vol. 83). Available from: https://books.google.com.eg/books?id=EQteuWJ3058C

33. Turner S, Pryer KM, Miao VPW, Palmer JD. Investigating Deep Phylogenetic Relationships among Cyanobacteria and Plastids by Small Subunit rRNA Sequence Analysis. J Eukaryot Microbiol [Internet]. 1999 Jul [cited 2019 Aug 13]; 46(4):327–38. Available from: https://onlinelibrary.wiley.com/doi/10.1111/j.1550-7408.1999.tb04612.x PMID: 10461381

34. Watts GS, Youens-Clark K, Slepian MJ, Wolk DM, Oshiro MM, Metzger GS, et al. {16S} {rRNA} gene sequencing on a benchtop sequencer: accuracy for identification of clinically important bacteria. J Appl Microbiol. 2017 Dec; 123(6):1584–96. https://doi.org/10.1111/jam.13590 PMID: 28940494

35. Heidelberg KB, Nelson WC, Holm JB, Eisenkolb N, Andrade K, Emerson JB. Characterization of eukaryotic microbial diversity in hypersaline {Lake} {Tyrrell}, {Australia}. Front Microbiol. 2013;4.

36. Wolf C, Kilias ES, Metfies K. Evaluating the potential of {18S} {rDNA} clone libraries to complement pyrosequencing data of marine protists with near full-length sequence information. Marine Biology Research. 2014 Sep; 10(8):771–80.

37. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. Appl Environ Microbiol. 2013; 79(17):5112–20. https://doi.org/10.1128/AEM.01043-13 PMID: 23793624

38. Sierra MA, Li Q, Pushalkar S, Paul B, Sandoval TA, Kamer AR, et al. The Influences of Bioinformatics Tools and Reference Databases in Analyzing the Human Oral Microbial Community. Genes (Basel) [Internet]. 2020 Aug 1 [cited 2021 Aug 10]; 11(8):1–12. Available from: /pmc/articles/PMC7465726/ https://doi.org/10.3390/genes11080878 PMID: 32756341

39. Mundra S, Kjønaas OJ, Morgado LN, Krabberød AK, Ransedokken Y, Kauserud H. Soil depth matters: Shift in composition and inter-kingdom co-occurrence patterns of microorganisms in forest soils. FEMS Microbiol Ecol. 2021 Feb 6; 97(3).

40. Lima IS, Guidelli EJ, Baffa— O, Vappou J, Bour P, Marquet F, et al. Utilization of different nitrogen sources for the growth of microalgae isolated from mangrove leaves in Banda Aceh—Indonesia. IOP Conf Ser Earth Environ Sci [Internet]. 2018 Nov 1 [cited 2022 Apr 28]; 207(1):012049. Available from: https://iopscience.iop.org/article/10.1088/1755-1315/207/1/012049

41. Kim G, Mujtaba G, Lee K, Kim G, Mujtaba G, Lee K. Effects of nitrogen sources on cell growth and biochemical composition of marine chlorophyte Tetraselmis sp. for lipid production. Algae [Internet]. 2016 Sep 30 [cited 2022 Apr 28]; 31(3):257–66. Available from: http://www.e-algae.org/journal/view.php?number=2789

42. Flynn K, Butler I. Nitrogen sources for the growth of marine microalgae: role of dissolved free amino acids. Mar Ecol Prog Ser. 1986; 34:281–304.

43. Gour RS, Bairagi M, Garlapati VK, Kant A. Enhanced microalgal lipid production with media engineering of potassium nitrate as a nitrogen source. https://doi.org/101080/2165597920171316440 [Internet]. 2017 Jan 1 [cited 2021 Aug 26]; 9(1):98–107. Available from: https://www.tandfonline.com/doi/abs/10.1080/21655979.2017.1316440 PMID: 28471319

44. Rana MS, Prajapati SK. Resolving the dilemma of iron bioavailability to microalgae for commercial sustenance. Algal Res [Internet]. 2021 Nov 1 [cited 2021 Aug 27]; 59:102458. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2211926421002770

45. Botebol H, Sutak R, Scheiber IF, Blaiseau PL, Bouget FY, Camadro JM, et al. Different iron sources to study the physiology and biochemistry of iron metabolism in marine micro-algae. BioMetals [Internet]. 2014 Feb 27 [cited 2021 Aug 27]; 27(1):75–88. Available from: http://link.springer.com/10.1007/s10534-013-9688-1 PMID: 24281777

46. Juneja A, Ceballos R, Murthy G. Effects of Environmental Factors and Nutrient Availability on the Biochemical Composition of Algae for Biofuels Production: A Review. Energies (Basel) [Internet]. 2013 Sep 3 [cited 2019 Nov 20]; 6(9):4607–38. Available from: http://www.mdpi.com/1996-1073/6/9/4607

47. Yang Z, Pei H, Han F, Wang Y, Hou Q, Chen Y. Effects of air bubble size on algal growth rate and lipid accumulation using fine-pore diffuser photobioreactors. Algal Res [Internet]. 2018 Jun [cited 2019 Dec 9]; 32:293–9. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2211926417310019

48. Berge T, Daugbjerg N, Hansen PJ. Isolation and cultivation of microalgae select for low growth rate and tolerance to high {pH}. Harmful Algae [Internet]. 2012 Dec [cited 2019 Dec 9]; 20:101–10. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1568988312001436

49. Vadlamani A, Viamajala S, Pendyala B, Varanasi S. Cultivation of Microalgae at Extreme Alkaline pH Conditions: A Novel Approach for Biofuel Production. ACS Sustain Chem Eng [Internet]. 2017 Aug 7 [cited 2021 Aug 10]; 5(8):7284–94. Available from: https://pubs.acs.org/doi/full/10.1021/acssuschemeng.7b01534

50. Nolan MP, Cardinale BJ. Species diversity of resident green algae slows the establishment and proliferation of the cyanobacterium {Microcystis} aeruginosa. Limnologica [Internet]. 2019 Jan [cited 2019 Aug 15]; 74:23–7. Available from: http://www.sciencedirect.com/science/article/pii/S0075951118301105

51. Robla J, García-Hierrro J, Alguacil FJ, Dittami SM, Marie D, Villa E, et al. Determination of the efficiency of filtration of cultures from microalgae and bacteria using hollow fiber filters. Environ Sci (Camb) [Internet]. 2021 Jul 1 [cited 2022 May 8]; 7(7):1230–9. Available from: https://pubs.rsc.org/en/content/articlehtml/2021/ew/d0ew00927j

52. Dzhembekova N, Rubino F, Nagai S, Zlateva I, Slabakova N, Ivanova P, et al. Comparative analysis of morphological and molecular approaches integrated into the study of the dinoflagellate biodiversity within the recently deposited Black Sea sediments–benefits and drawbacks. Biodivers Data J [Internet]. 2020 [cited 2022 May 20];8. Available from: /pmc/articles/PMC7447646/

53. Djurhuus A, Port J, Closek CJ, Yamahara KM, Romero-Maraccini O, Walz KR, et al. Evaluation of filtration and DNA extraction methods for environmental DNA biodiversity assessments across multiple trophic levels. Front Mar Sci. 2017 Oct 13; 4(OCT):314.

54. Muñoz-Colmenero M, Sánchez A, Correa B, Figueiras FG, Garrido JL, Sotelo CG. Evaluation of DNA Extraction Methods and Bioinformatic Pipelines for Marine Nano- and Pico-Eukaryotic Plankton Analysis. Front Mar Sci. 2021 Jan 14; 7:1220.

55. Thomas T, Gilbert J, Meyer F. Metagenomics—a guide from sampling to data analysis. Microb Inform Exp [Internet]. 2012 Dec [cited 2023 Apr 3]; 2(1):3. Available from: /pmc/articles/PMC3351745/ https://doi.org/10.1186/2042-5783-2-3 PMID: 22587947

56. Brandt MI, Pradillon F, Trouche B, Henry N, Liautard-Haag C, Cambon-Bonavita MA, et al. Evaluating sediment and water sampling methods for the estimation of deep-sea biodiversity using environmental DNA. Scientific Reports 2021 11:1 [Internet]. 2021 Apr 12 [cited 2022 May 4]; 11(1):1–14. Available from: https://www.nature.com/articles/s41598-021-86396-8 https://doi.org/10.1038/s41598-021-86396-8 PMID: 33846371

57. Harrison JB, Sunday JM, Rogers SM. Predicting the fate of eDNA in the environment and implications for studying biodiversity. Proceedings of the Royal Society B [Internet]. 2019 Nov 20 [cited 2022 May 3]; 286(1915). Available from: https://royalsocietypublishing.org/doi/full/10.1098/rspb.2019.1409 PMID: 31744434

58. Clare EL, Economou CK, Faulkes CG, Gilbert JD, Bennett F, Drinkwater R, et al. eDNAir: Proof of concept that animal DNA can be collected from air sampling. PeerJ [Internet]. 2021 Mar 31 [cited 2022 May 3]; 9:e11030. Available from: https://peerj.com/articles/11030 https://doi.org/10.7717/peerj.11030 PMID: 33850648

59. Manoylov KM. Taxonomic identification of algae (morphological and molecular): species concepts, methodologies, and their implications for ecological bioassessment. J Phycol [Internet]. 2014 Jun 1 [cited 2023 Jan 14]; 50(3):409–24. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/jpy.12183 PMID: 26988316

60. Fawley MW, Fawley KP. Identification of Eukaryotic Microalgal Strains. J Appl Phycol [Internet]. 2020 Oct 1 [cited 2023 Jan 14]; 32(5):2699. Available from: /pmc/articles/PMC7853647/ https://doi.org/10.1007/s10811-020-02190-5 PMID: 33542589

61. Shubert E, Gärtner G. Nonmotile Coccoid and Colonial Green Algae. In: Freshwater Algae of North America [Internet]. Elsevier; 2015 [cited 2019 Dec 9]. p. 315–73. Available from: https://linkinghub.elsevier.com/retrieve/pii/B9780123858764000074

62. Sassenhagen I, Wilken S, Godhe A, Rengefors K. Phenotypic plasticity and differentiation in an invasive freshwater microalga. Harmful Algae [Internet]. 2015 Jan [cited 2019 Dec 4]; 41:38–45. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1568988314001966

63. Hugenholtz P, Chuvochina M, Oren A, Parks DH, Soo RM. Prokaryotic taxonomy and nomenclature in the age of big sequence data. The ISME Journal 2021 15:7 [Internet]. 2021 Apr 6 [cited 2021 Aug 10]; 15(7):1879–92. Available from: https://www.nature.com/articles/s41396-021-00941-x https://doi.org/10.1038/s41396-021-00941-x PMID: 33824426

64. Federhen S. The NCBI Taxonomy database. Nucleic Acids Res [Internet]. 2012 Jan; 40(Database issue):D136–D143. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3245000/ https://doi.org/10.1093/nar/gkr1178 PMID: 22139910

65. Chrysosporum bergii CHAB1000 [Internet]. [cited 2022 May 20]. Available from: https://www.uniprot.org/taxonomy/929837

66. Kokociński M, Soininen J. New insights into the distribution of alien cyanobacterium Chrysosporum bergii (Nostocales, Cyanobacteria). Phycological Res [Internet]. 2019 Jul 1 [cited 2022 May 20]; 67 (3):208–14. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/pre.12373

67. Schoch CL, Ciufo S, Domrachev M, Hotton CL, Kannan S, Khovanskaya R, et al. NCBI Taxonomy: A comprehensive update on curation, resources and tools. Database. 2020;2020. https://doi.org/10.1093/database/baaa062 PMID: 32761142

68. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil PA, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nature Biotechnology 2018 36:10 [Internet]. 2018 Aug 27 [cited 2021 Aug 10]; 36(10):996–1004. Available from: https://www.nature.com/articles/nbt.4229 https://doi.org/10.1038/nbt.4229 PMID: 30148503

69. Hernández-Pérez A, Zamora-Briseño JA, Söderhäll K, Söderhäll I. Gut microbiome alterations in the crustacean Pacifastacus leniusculus exposed to environmental concentrations of antibiotics and effects on susceptibility to bacteria challenges. Dev Comp Immunol. 2021 Jun 24;104181. https://doi.org/10.1016/j.dci.2021.104181 PMID: 34175331

70. Wang H, Banerjee N, Liang Y, Wang G, Hoffman KL, Khan MF. Gut microbiome-host interactions in driving environmental pollutant trichloroethene-mediated autoimmunity. Toxicol Appl Pharmacol. 2021 Aug 1; 424:115597. https://doi.org/10.1016/j.taap.2021.115597 PMID: 34051218

71. Ramakodi MP. Effect of Amplicon Sequencing Depth in Environmental Microbiome Research. Current Microbiology 2021 78:3 [Internet]. 2021 Feb 3 [cited 2021 Aug 10]; 78(3):1026–33. Available from: https://link.springer.com/article/10.1007/s00284-021-02345-8 PMID: 33537885

72. Alvarenga DO, Fiore MF, Varani AM. A metagenomic approach to cyanobacterial genomics. Front Microbiol. 2017 May 9; 8(MAY):809. https://doi.org/10.3389/fmicb.2017.00809 PMID: 28536564

73. RNAcentral: The non-coding RNA sequence database [Internet]. 2021 [cited 2021 May 16]. Available from: https://rnacentral.org/

74. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res [Internet]. 2013 Jan 1 [cited 2021 May 16]; 41(D1). Available from: https://pubmed.ncbi.nlm.nih.gov/23193283/ https://doi.org/10.1093/nar/gks1219 PMID: 23193283

75. Eiler A, Drakare S, Bertilsson S, Pernthaler J, Peura S, Rofner C, et al. Unveiling Distribution Patterns of Freshwater Phytoplankton by a Next Generation Sequencing Based Approach. PLoS One [Internet]. 2013 Jan 29 [cited 2020 Jun 26]; 8(1). Available from: https://pubmed.ncbi.nlm.nih.gov/23349714/ https://doi.org/10.1371/journal.pone.0053516 PMID: 23349714

76. Pushpakumara BLDU, Tandon K, Willis A, Verbruggen H. Unravelling microalgal-bacterial interactions in aquatic ecosystems through 16S rRNA gene-based co-occurrence networks. Scientific Reports 2023 13:1 [Internet]. 2023 Feb 16 [cited 2023 Feb 27]; 13(1):1–10. Available from: https://www.nature.com/articles/s41598-023-27816-9 https://doi.org/10.1038/s41598-023-27816-9 PMID: 36797257