

Metagenomic Analysis Reveals Microbial Communities in Lake Qarun – Egypt

Mahmoud Mohamed Yosef^{1*}, Ali Mohamed Elrefy¹, Fawzy Ali Elfeky¹

¹ Biotechnology Department, Faculty of Agriculture, Al-Azhar University, Cairo, Egypt

* Corresponding author's e-mail: myiusef@azhar.edu.eg

ABSTRACT

Characterization of water quality in Lake Qarun indicated that the water is very poor for irrigation and aquatic life. At the same time, the bacterial community was represented mainly by five bacterial phyla with different proportions: *Firmicutes* (53%), *Proteobacteria* (33%), *Bacteroidetes* (7%), *Actinobacteria* (5%) and *Thermi* (1%). Furthermore, metagenomes prediction of bacterial communities using PICRUSt indicated important functional gene families associated with metabolism, environmental information, genetic information processing, and cellular processes. It is worth noting that Benzoate degradation had the highest average relative abundance, followed by aminobenzoate degradation among 18 individual KEGG pathways from xenobiotics biodegradation and metabolism which showed higher relative abundance. The obtained data indicate that a different source of pollution in Qarun Lakes has an impact on the bacterial community's structure, as well as the biota and is expected to cause health problems.

Keywords: bacterial community, Ion Torrent, Qarun Lake, 16S rRNA high throughput.

INTRODUCTION

Lake Qarun (Moeris) is a historic lake that has supported human culture for over 7,000 years (El-Shabrawy et al. 2014). The lake is currently experiencing severe water pollution. As a result of uncontrolled disposal practices of solid and liquid household and industrial waste, as well as pesticide contamination and lack of long-term sewage treatment (Shaaban, Hanafi, and Ibrahim 2016). The Egyptian Company (EMISAL), located on the Lake's southern shore, causes the water to be concentrated up to ten times its natural salinity (A. Mageed 1998). The increase in salinity is also due to such factors as high temperatures and changes in water evaporation rate over the last few hundred years (Al-Afify et al., 2019). Furthermore, increasing contamination of Lake Qarun water resources, as well as the resulting effects on the aquatic environment and human health, is a large source of concern. Therefore, the study of microbial communities in Lake Qarun in Egypt

is critical because wailed life, aquatic life and humans all rely on water for their survival. Microbiome is the study of microbial communities collected directly from their natural environment without prior culture. Because of recent breakthroughs in Next Generation Sequencing (NGS) technologies, there has been tremendous development in this field of research (Schuster 2008). Because over 99.8% of microorganisms in some situations are unable to be cultivated (Shah et al. 2011) The development of amplicon and shotgun genome sequencing technologies allowed for a better knowledge of the microbiota-host connection. This study provides important insight into the bacterial community structure under the influence of various types of pollution, using 16S rRNA high throughput amplicon sequencing to investigate the bacterial abundance and diversity in Lake Qarun. Thus, bioinformatic analytic techniques were employed to better understand the bacterial community function in the lake under diverse sources of contamination.

MATERIALS AND METHODS

Sampling site, collection, and filtration

The water samples were taken from the Qarun Lake, which is located between 30° 24' and 30° 49' E longitudes and 29°24' and 29°33' N latitudes. It is around 27 kilometers north of the Al Fayoum City and 80 kilometers southwest of Cairo.(Al-Afify *et al.* 2019). The water samples were collected in 500 mL sterile screw-capped bottles from three-point stations along the lake as depicted on the map (Fig. 1). They included:

- Station 1: From the front of the Shakshuk village,
- Station 2: From the front of the Khak village and
- Station 3: From the front of the Senours vilage (Shaaban, Hanafi, and Ibrahim 2016).

The sample was collected into an ice box and transferred into the laboratory within a few hours. Three sample were mixed and pooled in sterile screw-caped glass bottles, then filtered through cellulose filter paper (Whatman paper circles, diam. 70 mm, Grade 1, size 11 μ m) and the mixture was then filtered via (0.22 μ m nitrocellulose). The filtration process material was supplied to 50 mL sterilized plastic tube and stored at -20°C (Cabello-Yeves *et al.* 2020).

Analysis of physicochemical properties

The physical and chemical characteristics were estimated at Agriculture Research Center (A.R.C) in Soil, Water & Environment Research Institute laboratory. The water sample was collected into ice box and pooled together in sterile screw-caped glass bottles, then transferred into laboratory in 24 hours. Electrical conductivity (EC) and pH were measured using an EC and pH meter (Model AD31-Adwa CO., Hungary). The methods of water sample analysis including dissolved oxygen, major cations and anions, as well as the concentration of heavy metal elements, such as lead (Pb), copper (Cu), cadmium (Cd), nickel (NI) and zinc (Zn), were determined using the conventional method (Ibrahim and Ramzy 2013).

High throughput 16S rRNA amplification and sequencing

Genomic DNA was extracted with DNeasy Blood & Tissue Kit (Qiagen, Inc.) from the filtered water sample on a cellulose-nitrocellulose filter paper, according to the manufacturer's guidelines (Mwaikono *et al.* 2016). Following quality and quantity checks, DNA extracted was utilized to amplify 16S rRNA gene using primer pair 338F, 907R and 1100F, which targets V3, V4-V5 and V7–9 respectively and PCR master mix (BioLabs, New England) (Fantini *et al.* 2015). High throughput



Figure 1. (a) The location of Lake Qarun in the southwest of Cairo and from the Nile Valley (<http://www.touregypt.net>) and (B, C and D) the distribution and location of each (point stations) of sampling sites on the lake.

16S rRNA amplicon sequencing was performed using 318 microchips of Ion Torrent PGM platform. According to Ion PGM Hi-Q View OT2 Kit User Guide (Thermo Fisher Scientific, USA) at Genetic Engineering and Biotechnology Research Institute, Sadat City, Al Buhayrah, Egypt.

Microbial data analysis

QIIME package release QIIME 1.9.1 64-bit was used to process and analyze all reads (Caporaso et al. 2010). Any low quality and chimeric sequences were removed from the analysis. Using the default settings, sequences were grouped into operational taxonomic units (OTUs), by the open-reference OTU clustering technique (with 97% similarity). OTU representative sequences assigned to taxonomy using both RDP Classifier and Greengenes database (McDonald et al. 2012), (Wang et al. 2007). Alpha diversity including Simpson index values, ACE estimator, Shannon index, Chao1 estimator and were estimated as indicators for alpha diversity.

Metagenomic prediction with PICRUSt

Bacterial function and metabolic pathways were predicted using the PICRUSt software (Langille et al. 2013). Using the Greengenes database's closed-reference OTU collection, an OTU table with 97% nucleotide identity was created in QIIME (McDonald et al. 2012). Normalization was conducted using the number of 16S rRNA copies per OTU, and metagenomes were inferred using the normalized OTU table (Jiménez et al. 2014). The PICRUSt software was also utilized to predict xenobiotic degradation and metabolism functions (Zhang et al. 2020). Raw data of high-throughput of 16S rRNA sequencing (fastq file) are available from the NCBI's Sequence Read Archive under accession (PRJNA789242).

RESULTS

Physical and chemical parameters of water sample

Physicochemical parameters of the water sample were determined (Table 1). Water Electrical conductivity EC in Lake Qarun was (45 mS/cm). The pH of natural waters is a significant component in their chemical and biological processes, in the Qarun Lake water, the pH value was (7.6), which usually lies on the alkaline side. Total Dissolved Solid was (36 g/L) where the salinity is comparable to that of sea water, while Total Dissolved Oxygen was (7.29 mg/L). Moreover, bicarbonate HCO_3^- , Chlorine Cl^- and sulphate SO_4^{2-} were measured as the major components of water alkalinity and their concentrations amounted to 0.35, 11.36 and 9.8 g/L respectively, while Calcium, Ca_2^+ Sodium Na^+ , Potassium, K^+ , and Magnesium, Mg_2^+ were measured as major cations and their concentrations equaled 8.9, 0.16, 0.72 and 1.24 g/L respectively. The concentration level of nutrient salts NH_4^+ and NO_3^- were 3.5 and 3.7 mg/L. The concentrations of many heavy metals were presented in Lake Qarun (Table 1).

Diversities and composition of bacterial communities on Lake Qarun

The total number of all input sequences was 185,338. After trimming, screening, and removing chimeras and single pieces, 48,942 were removed, 136,396 high quality-filtered sequences with the median sequence length: 239 bp (Fig. 2), 473 operational taxonomic units (OTUs) were assigned. The alpha diversity indices and richness were estimated as shown in Table 2. OTU numbers and alpha indices, including Shannon, Simpson, Chao 1, Fisher's

Table 1. Physical and chemical characteristics of Lake Qarun water

Physicochemical properties											
Water sample	pH	DO (mg/l)	TDS (g/L)	EC (mS/cm)	Major anions			Major cations			
					HCO_3^- (g/L)	Cl^- (g/L)	SO_4 (g/L)	Ca_2^+ (g/L)	Mg_2^+ (g/L)	Na^+ (g/L)	K^+ (g/L)
	7.5	7.29	36	45	0.35	11.36	9.8	0.72	1.24	8.9	0.16

Heavy metal elements (mg/L)												
Water sample	NH_4^+ (mg/L)	NO_3^- (mg/L)	B (mg/L)	Cu (mg/L)	Zn (mg/L)	Pb (mg/L)	Cd (mg/L)	P (mg/L)	Fe (mg/L)	Mn (mg/L)	Ni (mg/L)	Co (mg/L)
	3.5	3.7	0.015	0.013	0.050	0.20	1.99	0.010	0.133	0.065	1.51	0.74

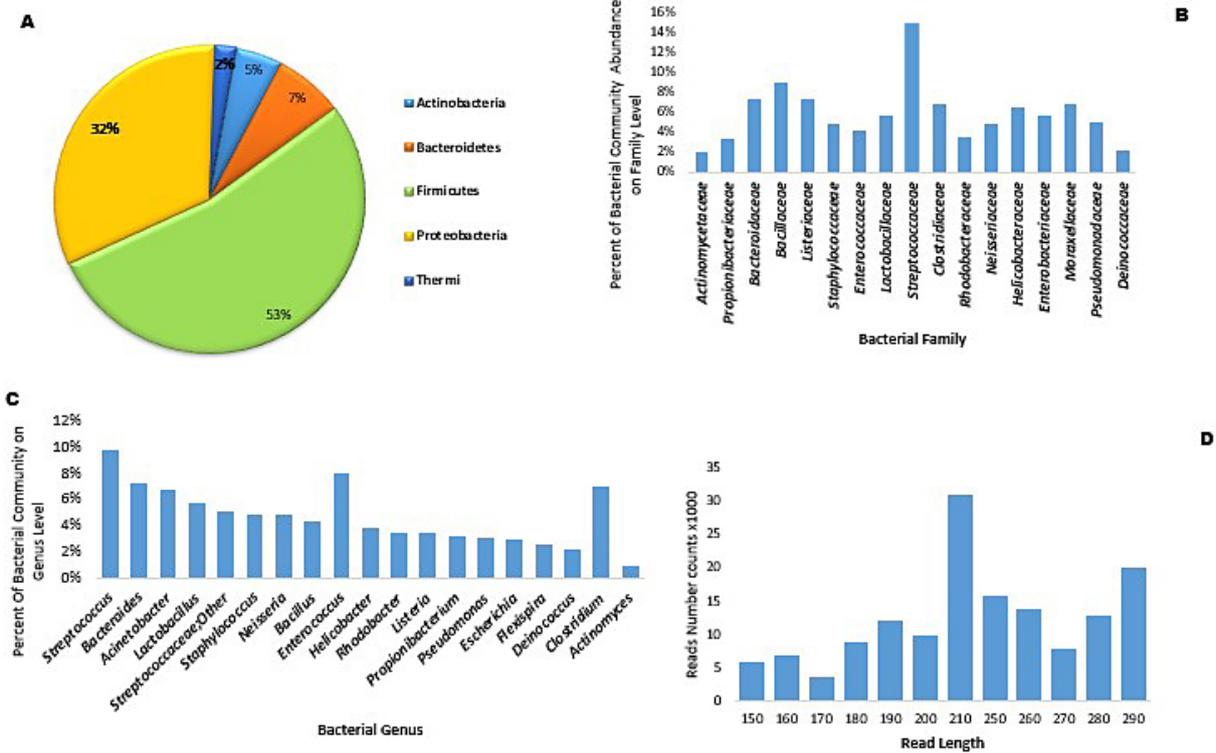


Figure 2. Relative abundance of bacterial community (based on 16S rRNA gene) in Qarun lake water at different taxonomic levels (A) Relative abundance of bacterial phylum (B) Relative abundance of bacterial Families (C) Relative abundance of the bacterial genus.(D) Histogram explains number of reads corresponding to the length of sequences for each reads. The average of sequence length was: 239 pb

and phylogenetic diversity which indicate that sequencing had high diversity. Five bacterial phyla were detected with different proportions *Firmicutes* (53%), *Proteobacteria* (33%), *Bacteroidetes* (7%), *Actinobacteria* (5%) and *Thermi* (1%). In order to better explain the relative abundance of bacterial communities was analyzed at different taxonomic levels family and genus (Fig. 2). At family level *Streptococcaceae* (14%), *Bacteroidaceae* (7%), *Listeriaceae* (7%), *Clostridiaceae* (7%), *Moraxellaceae* (7%), *Helicobacteraceae* (6%), *Lactobacillaceae* (6%), and *Enterobacteriaceae* (6%), were the most abundant. Unclassified bacteria accounted for about 5% of all sequences at family level. *Streptococcus* was higher abundant genus with (10%) than other sequences at the genus level followed *Bacteroides* (7%) and *Lactobacillus* (6%) (Fig. 2).

PICRUSt functional predictive analysis

The use of the PICRUSt tool led to the discovery of important functional gene families associated with metabolism, environmental information, genetic information processing, and cellular processes. In all sequences, metabolism had the highest relative abundance, followed by genetic information processing (Fig. 3). At level 2, many sequences have been referred to amino acid metabolism and carbohydrate metabolism at (Kyoto Encyclopedia of Genes and Genomes). In turn, at level 3 the efforts were concentrated on the functional genes involved in xenobiotic biodegradation (Fig. 4). Most aquatic life organisms including bacteria, are exposed to these xenobiotics, because the lake is afflicted with severe water pollution as a result of numerous sources of industrial and agricultural pollution, as well as sewage issues (Shaaban et al., 2016). Consequently, 18 KEGG

Table 2. statistical table explain quality reads, observed OTU numbers, Shannon, Simpson, Chao 1, Fisher’s alpha indices and phylogenetic diversity

Sample	Sequence	OTSu	Chaol	Simpson	Shannon	Fisher_alpha	PD_whole_tree
Qarun lake water	136,396	473	473	0.991	7.387	61.515	7.480

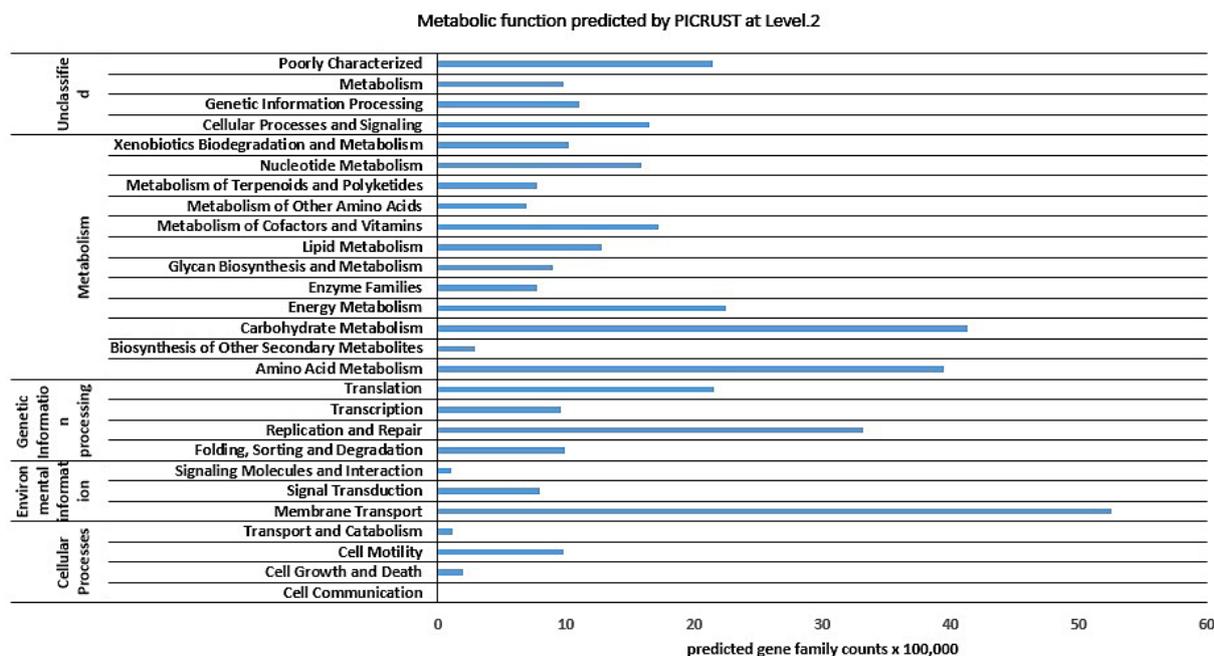


Figure 3. Bacterial community functions predicted by PICRUST at level 3 which related to xenobiotics degradation and metabolism

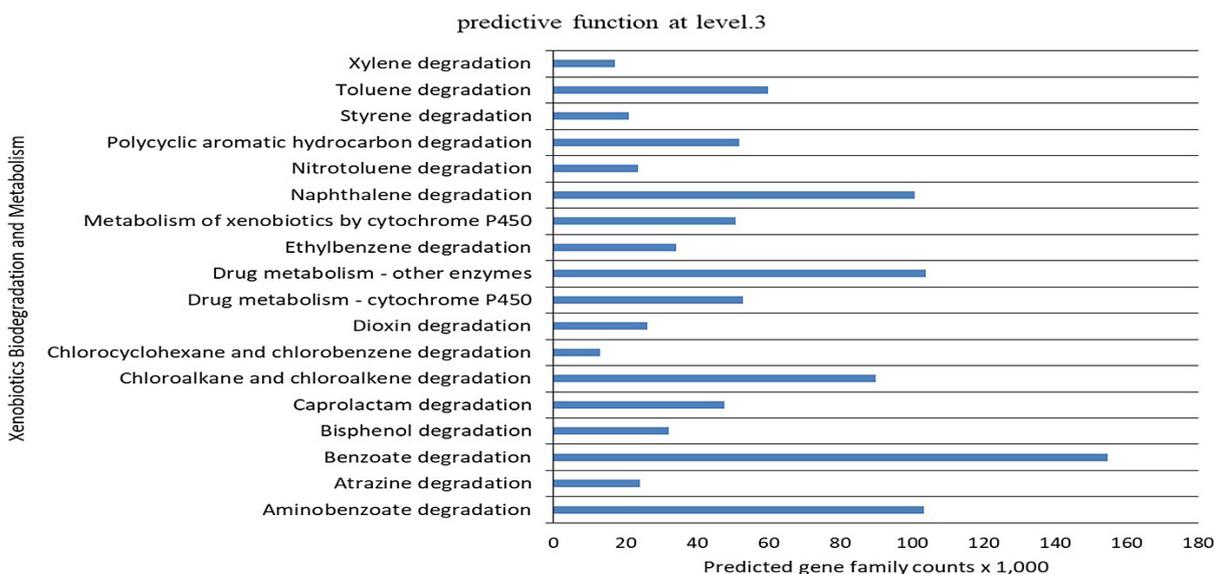


Figure 4. Bacterial community functions predicted by PICRUST at level 3 which related to xenobiotics degradation and metabolism

pathways related to xenobiotic biodegradation and metabolism were chosen and exhibited higher relative abundance. The average relative abundance of benzoate degradation was the highest among them, followed by aminobenzoate degradation.

DISCUSSION

The rising pollution levels in the Lake Qarun water, as well as the consequences for the aquatic ecosystem and human health, have long been a

focus of numerous researchers. Many prior investigations on the physical and chemical features of the Lake Qarun water quality were conducted (Abdel Wahed et al., 2015); (Al-Afify et al., 2019); (Ali et al. 2010); (Authman et al., 2007); (Gupta and Abd El-Hamid 2003); (Mansour and Sidky 2003). They were all in agreement that the Lake Qarun water was rated (Poor) for irrigation standards and (Very Poor) for aquatic life guidelines, according to the overall average value of the water quality index (WQI). Furthermore, the metal indices revealed that the Qarun Lake water

is contaminated to varying degrees for the aquatic life use due to the massive amount of waste that is discharged into the lake (Goher et al. 2018). The physical and chemical parameters in this study agree with the previous studies with slight variations. A variety of reasons could be to blame for the disparity in results. The initial sampling site variance is influenced by the distance from the Lake's water intake and wastewater outflow locations, as well as the sampling method used and the seasons' time difference (Dang, Speck, and Benkendorff 2012). Bacterial populations are essential for assuring and sustaining environmental and ecological processes in lake environments (Araya et al. 2003). The result in this research revealed the highest abundance values for members of the Gram-positive phylum, *Firmicutes* the most dominant phylum while covered (53%) of the total number of reads in the water community. This result was consistent with the result of (Shaban et al., 2016), who investigated the distribution of microorganisms in Lake Qarun and found that distinct members of the *Firmicutes* bacteria were identified as indicators of their abundance. Many members of the *Firmicutes* are notable for their ability to break down even the most resistant organic substances. The main classes of *Firmicutes* were Bacilli, which had the largest relative abundance in the water of Lake Qarun. Bacilli in freshwater are thought to be signs of human faeces contamination (Zinger et al., 2011). In the bacterial community, the *Streptococcaceae* family is the most abundant, *Streptococcus sp.* is extremely pathogenic, affecting a wide range of fish species worldwide. Their prevalence and severity are influenced by various environmental conditions, including ammonia levels, dissolved oxygen levels and water temperatures. Furthermore, these conditions allow a wide range of anaerobic microorganisms to grow well (El-Noby et al. 2021). Besides the predominating *Firmicutes*, high proportions of Proteobacteria were found, which covered (33%) of the total number of reads in the water communities. *Proteobacteria* contain many members of the Proteobacteria families such as *Rhodobacteraceae* (Alphaproteobacteria), *Neisseriaceae* (Betaproteobacteria) and *Pseudomonadaceae* (Gammaproteobacteria) that are halotolerant bacterial population playing a significant role in the bioremediation of anthropogenic compounds and in biological reactor systems, such as wastewater treatment plants (Köchling et al. 2017). These bacteria are noted

for their ubiquity and metabolic flexibility, which includes the capacity to withstand harsh environments, use a variety of carbon molecules, and live both aerobic and anaerobic lifestyles (Paul et al. 2016). Additionally, *Bacteroidetes* were detected as a major group in the bacterial communities of the Lake Qarun water. GBac (General Bacteroidales Marker) indicates extensive faecal contamination in the environment (Kirs et al. 2011). The high abundance of *Bacteroidales* and other bacteria indicates widespread faecal pollution in Lake Qarun as a result of untreated wastewater discharge points throughout the lake's lengthy coast. Bacteroidetes in the environment are hypothesized to specialize in degrading complex organic substances in the biosphere, particularly polysaccharides and proteins. They may utilize a wide range of biopolymers as carbon and energy sources, such as plant, algal, and animal molecules (Thomas et al. 2011).

CONCLUSIONS

On the basis of 16S rRNA high throughput amplicon sequencing on the Ion torrent platform, the first analysis of the composition and organization of bacterial communities in Lake Qarun was reported. According to the obtained findings, environmental factors have a substantial impact on the composition of bacterial communities and their ability to endure harsh settings. Furthermore, the taxonomic assignment of the NGS sequence validated faecal contamination and provided the information on pathogenic bacteria for aquatic life. It is recommended that liquid waste disposal in rivers, lakes, or seas be avoided, and effective planning must be implemented.

Acknowledgments

The authors acknowledge the researchers at the Genetic Engineering and Biotechnology Research Institute in Sadat City and Beheira, as well as Dr. Walid S. Mohamed and Dr. Mohammed Ezz Al Rijal, for their contributions to the study.

REFERENCES

1. Mageed A.A. 1998. Distribution and Salinity Ranges of Zooplankton Organisms at El-Fayoum Depression (El-Fayoum-Egypt). Egyptian Journal of Aquatic Biology and Fisheries, 2(2), 51–71.

2. Wahed S.M.A., Mohamed E.A., Wolkersdorfer C., El-Sayed M.I., M'nif A., Sillanpää M. 2015. Assessment of Water Quality in Surface Waters of the Fayoum Watershed, Egypt. *Environmental Earth Sciences*, 74(2), 1765–1783.
3. Al-Afify A.D.G., Tahoun U.M., Abdo M.H. 2019. Water Quality Index and Microbial Assessment of Lake Qarun, El-Batts and El-Wadi Drains, Fayoum Province, Egypt. *Egyptian Journal of Aquatic Biology and Fisheries*, 23(1), 341–357.
4. Ali F.K., El-Shafai S.A., Samhan F.A., Khalil W.K.B. 2010. Effect of Water Pollution on Expression of Immune Response Genes of Solea Aegyptiaca in Lake Qarun. *African Journal of Biotechnology*, 7(10), 1418–1425.
5. Araya R., Katsuji T., Tatsuya T., Nobuyasu Y., Masao N. 2003. Bacterial Activity and Community Composition in Stream Water and Biofilm from an Urban River Determined by Fluorescent in Situ Hybridization and DGGE Analysis. *FEMS Microbiology Ecology*, 43(1), 111–119.
6. Cabello-Yeves P.J., Zenskaya T.I., Zakharenko A.S., Sakirko M.V., Ivanov V.G., Ghai R., Rodriguez-Valera F. 2020. Microbiome of the Deep Lake Baikal, a Unique Oxic Bathypelagic Habitat. *Limnology and Oceanography*, 65(7), 1471–1488.
7. Caporaso J.G., Kuczynski J., Stombaugh J., Bittinger K., Bushman F.D., Costello E.K., Fierer N., et al. 2010. QIIME Allows Analysis of High-Throughput Community Sequencing Data. *Nature Methods*, 7(5), 335–336.
8. Dang V.T., Speck P., Benkendorff K. 2012. Influence of Elevated Temperatures on the Immune Response of Abalone, *Haliotis Rubra*. *Fish & Shellfish Immunology*, 32(5), 732–740.
9. El-Noby G.A., Hassanin M., El-Hady M., Aboshabana S. 2021. *Streptococcus*: A Review Article on an Emerging Pathogen of Farmed Fishes. *Egyptian Journal of Aquatic Biology and Fisheries*, 25(1), 123–139.
10. El-Shabrawy G., Gohar M., Germoush M., Anufrieva E., Shadrin N. 2014. Long-Term Changes in the Saline Lake Qarun (Egypt). *Acta Geologica Sinica*, 88, 180–182.
11. Fantini E., Gianese G., Giuliano G., Fiore A. 2015. Bacterial Metabarcoding by 16S rRNA Gene Ion Torrent Amplicon Sequencing. *Methods in Molecular Biology*, 1231, 77–90.
12. Goher M.E., El-Rouby W.M.A., El-Dek S.I., El-Sayed S.M., Noaemy S.G. 2018. Water Quality Assessment of Qarun Lake and Heavy Metals Decontamination from Its Drains Using Nanocomposites. *Materials Science and Engineering*, 464(1).
13. Gupta G., El-Hamid Z.A. 2003. Water Quality of Lake Qarun, Egypt. *International Journal of Environmental Studies*, 60(6), 651–657.
14. Lubna I., Ramzy E.M. 2013. Water Quality and Its Impact on *Tilapia Zilli* (Case Study) Qarun Lake-Egypt. *International Water Technology Journal*, 3(2).
15. Kirs M., Harwood V.J., Fidler A.E., Gillespie P.A., Fyfe W.R., Blackwood A.D., Cornelisen C.D. 2011. Source Tracking Faecal Contamination in an Urbanised and a Rural Waterway in the Nelson-Tasman Region, New Zealand. *New Zealand Journal of Marine and Freshwater Research*, 45(1), 43–58.
16. Köchling T., Sanz J.L., Galdino L., Florencio L., Kato M.T. 2017. Impact of Pollution on the Microbial Diversity of a Tropical River in an Urbanized Region of Northeastern Brazil. *International Microbiology*, 20(1), 11–24.
17. Mansour S.A., Sidky M.M. 2003. *Ecotoxicological Studies. 6. The First Comparative Study between Lake Qarun and Wadi El-Rayan Wetland (Egypt), with Respect to Contamination of Their Major Components*. Food Chemistry. Elsevier, 82(2), 181–189.
18. McDonald D., Price M.N., Goodrich J., Nawrocki E.P., Desantis T.Z., Probst A., Andersen G.L., Knight R., Hugenholtz P. 2012. An Improved Greengenes Taxonomy with Explicit Ranks for Ecological and Evolutionary Analyses of Bacteria and Archaea. *ISME Journal*, 6(3), 610–618.
19. Authman M.M., Abbas H. 2007. Accumulation and Distribution of Copper and Zinc in Both Water and Some Vital Tissues of Two Fish Species (*Tilapia Zillii* and *Mugil Cephalus*) of Lake Qarun, Fayoum Province, Egypt. *Pakistan Journal of Biological Sciences: PJBS*, 10(13), 2106–2122.
20. Samson M.K., Maina S., Aswathy S., Schilling M., Kapur V., Gwakisa P. 2016. High-Throughput Sequencing of 16S rRNA Gene Reveals Substantial Bacterial Diversity on the Municipal Dumpsite. *BMC Microbiology*, 16(1), 1–12.
21. Dhiraj P., Kumbhare S.V., Mhatre S.S., Chowdhury S.P., Shetty S.A., Marathe N.P., Bhute S., Shouche Y.S. 2016. Exploration of Microbial Diversity and Community Structure of Lonar Lake: The Only Hypersaline Meteorite Crater Lake within Basalt Rock. *Frontiers in Microbiology*. Frontiers, 1553.
22. Schuster S.C. 2008. Next-Generation Sequencing Transforms Today's Biology. *Nature Methods*, 5(1), 16–18.
23. Shaaban M.T., Hanafi A.A.M., Ibrahim H.A.H. 2016. Distribution of Bacteria in Lake Qarun, AL Fayoum, Egypt (2014–2015) in Relation to Its Physical and Hydrochemical Characterization. *Journal of Bioscience and Applied Research*, 2(9), 601–615.
24. Neethu S., Tang H., Doak T.G., Ye Y. 2011. Comparing Bacterial Communities Inferred from 16S rRNA Gene Sequencing and Shotgun Metagenomics. *proc. Pacific Symposium on Biocomputing*, 165–176.
25. Thermo Fisher, Ion PGM Hi-Q View OT2 Kit User Guide (Pub. No. MAN0014579 Rev. C.0)
26. François T., Hehemann J.H., Rebuffet E., Czjzek M., Michel G. 2011. Environmental and Gut Bacteroidetes. *The Food Connection*. *Frontiers in Microbiology*, 2, 93.
27. Qiong W., Garrity G.M., Tiedje J.M., Cole J.R. 2007. Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Applied and Environmental Microbiology*, 73(16), 5261–5267.