Physicochemical Impacts on Bacterial Communities in Putrajaya Lake, Malaysia

Nurul Najwa Farihah Mat Lazim[1], Afiqah Mohamed[1], Zana Ruhaizat Zana Rudin[1], Fatimah Md Yusoff[2,3], Ikhsan Natrah[3] and Shahrizim Zulkifly[1]*

[1]*Department of Biology, Faculty of Science, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia*
[2]*International Institute of Aquaculture and Aquatic Sciences, Universiti Putra Malaysia, 71050 Port Dickson, Negeri Sembilan, Malaysia*
[3]*Department of Aquaculture, Faculty of Agriculture, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia*

ABSTRACT

This study determines the associations between the bacterial communities and water physicochemical parameters in Putrajaya Lake and Putrajaya Wetlands Park, Malaysia. Bacterial communities were assessed by metagenomics of the 16S rRNA gene from lake water input, central wetlands, and primary lake area. Water samples (n=18) were collected during two different periods: post-high rainfall events (samples collected in May) and dry periods (July). The data revealed that bacterial communities of the three sites were taxonomically distinct and associated with different environmental parameters. However, no significant differences were found between the wet and dry periods. Alpha diversity analyses revealed the highest index in May 2018 in the constructed wetlands (H'= 5.397) than those from water input or primary lake (p<0.05). Overall, 49 phyla, 147 classes, 284 orders, 471 families, 778 genera and 62 species of bacteria were identified. Verrumicrobia and Firmicutes showed a strong positive correlation with ammonia-nitrogen (r = 0.709). Actinobacteria and Cyanobacteria had a moderate positive correlation with nitrate with r value (r = 0.673) and (r = 0.647), respectively. In this study, the metagenomics of the 16S rRNA gene amplicon by Illumina MiSeq has successfully identified the bacterial community assemblage in Putrajaya Lake and wetlands. Bacterial composition was

Nurul Najwa Farihah Mat Lazim, Afiqah Mohamed, Zana Ruhaizat Zana Rudin, Fatimah Md Yusoff, Ikhsan Natrah and Shahrizim Zulkifly

associated with the availability of physicochemical properties of specific sites. The effectiveness of the engineered wetlands of Putrajaya in bioremediation was demonstrated by the marked decrease in certain nutrient concentrations from lake water input to the primary lake area.

Keywords: 16S rRNA, bacterial community, engineered wetlands, physicochemical, water quality

INTRODUCTION

Environmental concerns, which include climate change, habitat degradation, and water pollution, have directly or indirectly affected aquatic ecosystems' diversity and water quality. Putrajaya Lake is one of the prides of Putrajaya and is located in Malaysia's federal government administrative centre and has a total surface area of 400 ha. It has been designed to meet Class IIB of the Malaysian National Water Quality Standard (NWQS) and is suitable for body-contact recreational activities (Suratman et al., 2016). Putrajaya Holdings is entrusted to manage and maintain Putrajaya Lake by evaluating and monitoring the water quality according to the National Water Quality Standards, which have implemented various smart initiatives, such as low-carbon initiatives, with seven focal areas, in line with the Sustainable Development Goals of the United Nations (Majizat et al., 2016).

The Putrajaya Wetlands Park is an essential part of Putrajaya Lake, filtering the water from the primary lake from two main rivers: Chuau River and Bisa River. This engineered wetland covers 200 ha and was constructed as a natural filtration and bioremediation system of water input to the lake.

In Putrajaya Wetlands Park, several different plant communities have been planted to serve as a nutrient filtration system (Newton et al., 2011). Phytoremediation is a biological process in which selected plants remove water contaminants or pollutants. The filtration system involves the accumulation of contaminants by plant roots and provides a root zone and sediment habitat for bacteria and microbes that filter and decompose contaminants (Mohamad, 2012; Sabkie et al., 2020).

Rapid population growth and urbanisation lead to inevitable repercussions on the environment, such as releasing nutrient-rich runoffs into the lake, which have been cited as the main factors affecting reduced lake quality. These continuous changes in the seasonal and physicochemical parameters have undoubtedly influenced the state of water quality in Putrajaya Lake. Many indications of pollution were found within Putrajaya Lake due to the anthropogenic activities from the inlets or surface run-off points surrounding the lake (Asmat et al., 2018). Monitoring the nutrient removal for water quality improvement has been done at a pilot scale to simulate the essential roles of inhabiting macrophytes within the constructed wetlands (Vymazal, 2007). As far as the water quality monitoring assessment of freshwater lake is concerned, the Putrajaya Corporations (PjC) has developed the Putrajaya Lake Water Quality Standards (PLWQS) (Sharip et al., 2016). A previous study which aimed to design an

algorithm prediction for water quality changes has found that ammonia concentration plays a crucial role in determining the water quality level (Najah et al., 2021).

In addition to the current monitoring of the water quality parameters, there is a need to understand Putrajaya Lake in a more holistic approach, such as in biodiversity studies. Biodiversity is one of the critical components that determine the well-being of the lake, as it can be used as an indicator to assess the lake's health.

Previous studies in Putrajaya Lake only focused on assessing water quality and diversity of phytoplankton. Research on phytoplankton community structure as an ecosystem health indicator in the management of Putrajaya Lake and Putrajaya Wetlands Park indicated that there were 148 species from 77 genera identified during a sampling period from October 2009 until September 2010 (Jamal et al., 2014). Another study from November 2017 until January 2018 on phytoplankton diversity from three sites in Putrajaya Lake and Putrajaya Wetlands Park reported the discovery of 14 genera within eight classes (Sabkie et al., 2020). There was a higher density of phytoplankton recorded in Putrajaya Lake compared to the wetland.

The bacteria population in the lake forms a vital niche in nutrient cycling, decomposition and remediation. Bacteria are essential for all life in the ecosystem as they play a role in maintaining the structure, function, and sustainability of the ecosystem (Briones & Raskin, 2003). Information on bacterial diversity has the potential to aid water quality assessments. This information is currently limited to Putrajaya Lake.

Bacteria have a significant role in regenerating and mobilising the nutrients in freshwater food webs and are important in recycling the most naturally dynamic components in the ecosystem (Newton et al., 2011). Bacteria collectively are responsible for the movement of substances and conditions of water bodies (Pernthaler & Amann, 2005) as a result of their biomass creation and trophic coupling to eukaryotic predators as well as the foremost essential degraders and converters of an organic compound into inorganic material (Cotner & Biddanda, 2002).

The identification and classification of ubiquitous prokaryotes are widely investigated (Gupta et al., 2013). Bacteria can be identified based on morphology, biochemical tests, and culturing with selective media. However, not all environmental bacteria can be cultured under artificial laboratory conditions. The most current and practical approach to bacterial identification is through molecular methods using the Next Generation Sequencing (NGS) technique.

The next-generation sequencing (NGS) is one of the new technologies in the field of genomic analysis for DNA sequencing. It is a persuasive tool for demonstrating the diversity of various samples and studying metabolic pathways (Al-Sulaiman, 2012). The application of NGS leads to a high throughput accurate sequence read length and has enabled the investigation of a vast number of samples at a greater depth (Fadrosh et al., 2014).

Metagenomics studies used the direct genetic material of samples from the environment, bypassing the culturing process. The Illumina MiSeq platform can examine the community

composition of the clinical and environmental samples (Fadrosh et al., 2014). The 16S ribosomal RNA (rRNA) gene has been previously used for the assessment of bacteria variety and structure, including soil (Liles et al., 2003), marine (Sogin et al., 2006), and freshwater environments (Mueller-Spitz et al., 2009). The 16S rRNA gene is present in all bacteria, consisting of conserved regions to allow the design of a universal primer to amplify the 16S rRNA gene via polymerase chain reaction (PCR); incorporating the hypervariable regions (V1-V9) will be used for bacteria identification (Jo et al., 2016). The characterisation of bacteria community by using 16S rDNA as a molecular marker that targets hypervariable regions (V3, V5 or V6) for amplification allows profound sequencing (Staley et al., 2013), and V3-V4 regions have been used for the identification of bacterial genera from oligotrophic freshwater reservoirs using Illumina MiSeq platform (Iliev et al., 2017).

Putrajaya experiences an equatorial climate characterised by hot and humid weather all year round. Microclimate rainfall events and anthropogenic activities surrounding Putrajaya Lake are significant contributors to the lake's water quality, which include the release of effluents from the point and non-point sources, sewage from residential houses and commercial areas, the use of fertilisers for farming and rubbish littering. Identifying the bacteria community is important as bacteria can be an environmental bioindicator to detect changes in the lake's water quality. Many factors, which include global climate change, can threaten and shift the diversity of aquatic organisms in the water body. Water temperature has the potential to influence the changes in the hydrology of a water body and the habitat suitability of species.

In this study, Illumina MiSeq Sequencing was used to identify the bacterial diversity of Putrajaya Lake and Putrajaya Wetlands Park, targeting 16S rRNA V3-V4 hypervariable regions. Identifying the bacterial community will be a useful environmental indicator to detect changes in the lake's nutrient composition and water quality.

The objectives of this research were (1) to utilise the 16S ribosomal RNA (rRNA) gene for identifying bacteria species collected from water input, water after going through natural remediation of wetlands and primary lakes using next-generation sequencing technique, and (2) to determine the correlation between water body physicochemical properties and bacterial diversity. Here, we present the implementation of amplicon sequencing, which targeted the V3–V4 region of the 16S rRNA gene and the association of the physicochemical properties for studying bacterial community composition in Putrajaya Lake and Putrajaya Wetlands Park, Malaysia.

MATERIALS AND METHODS

Study Area and Sample Collection

Putrajaya Lake, located in the Federal Territory of Putrajaya, Malaysia, is an artificial freshwater lake sourced from the Chuau and Bisa River systems. Water sampling was

conducted in 2018 at three sites (Figure 1). The descriptions of the chosen sites are shown in Table 1.

Five hundred (500) mL water samples were collected in triplicate at each of the three sites using a Van Dorn sampler and collected in 500 mL PET sterilised water bottles. Surface water sampling (~1m) was done in selected months beginning November 2017 to August 2018. However, for this paper, only two months of data were selected, corresponding to

Table 1
Sampling sites

| Sites	Location	GPS Coordinate	Site Description
Site 1	The point of lake water input, Upper North 8 inlet (UN8; lake inlet)	2º59'10 "N, 101º412'26 "E	The main water inlet from Chuau River, the longest arm within the constructed wetland with diverse macrophyte species
Site 2	The Central Wetland (CW; engineered wetland)	2º57'1 "N, 101º41'37 "E	Final catchment area for bioremediated water before discharged into the primary lake area
Site 3	The open water at the Primary Lake f (PLf)	2º54'19.63 "N, 101º40'21.59 "E	Centre for water sports, recreational activities, and tourism
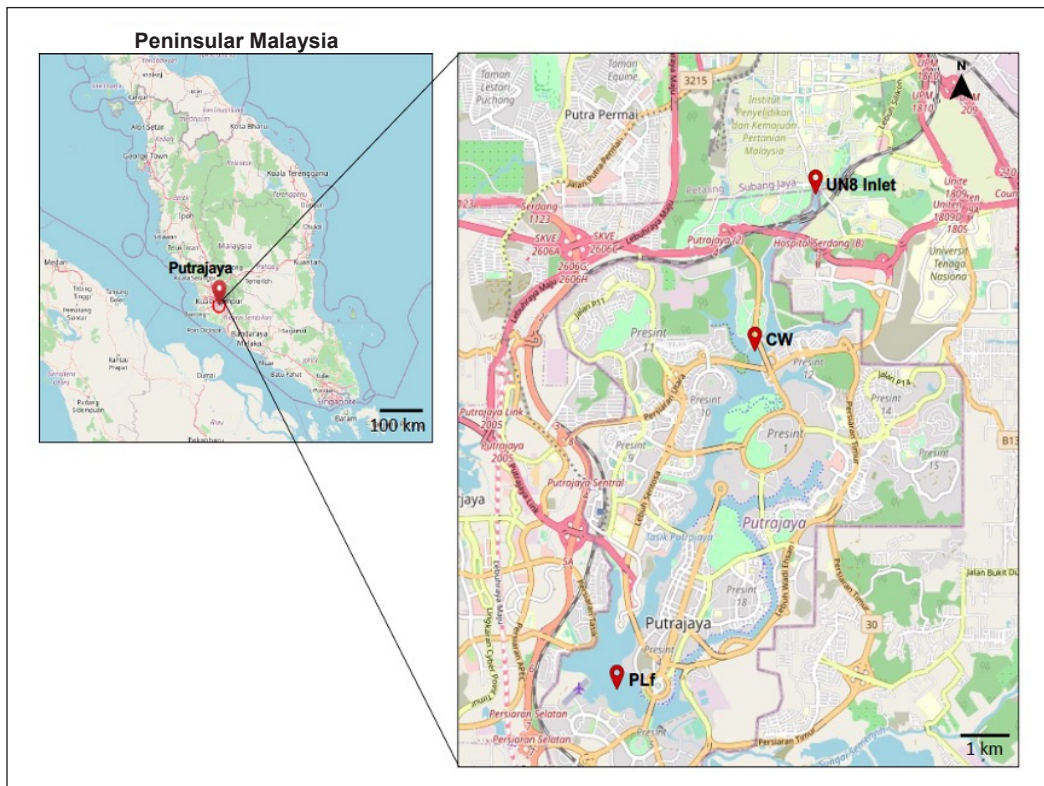
Figure 1. Three sampling sites in Putrajaya Lake and Putrajaya Wetlands Park, Site 1 (UN8: Upper North 8; lake inlet), Site 2 (CW: Central Wetland; engineered wetland), and Site 3 (PLf: Primary Lake f; open lake water)

the availability of metagenomics data and physicochemical correlations. The selection data was based primarily and limited on the samples' genomic quality and availability. In total, eighteen water samples were collected in May and July 2018 from three sources: Site 1: Upper North 8 (S1R1, S1R2, S1R3), Site 2: Central Wetland (S2R1, S2R2, S2R3), and Site 3: Primary Lake f (S3R1, S3R2, S3R3) were used in this study. Data in May 2018 was chosen to represent the wet period during the post-rainfall event, whereas data in July 2018 was chosen to represent the dry periods with less rainfall. The rainfall pattern data was provided by the Putrajaya Corporation (PjC) (Figure 2).

A total of eighteen samples collected represented three sources: Site 1 (S1R1, S1R2, S1R3), Site 2 (S2R1, S2R2, S2R3), and Site 3 (S3R1, S3R2, S3R3). All samples were immediately stored on ice in a cooler, transferred into the laboratory within 4 hours of collection, and stored at 4°C before further processing.
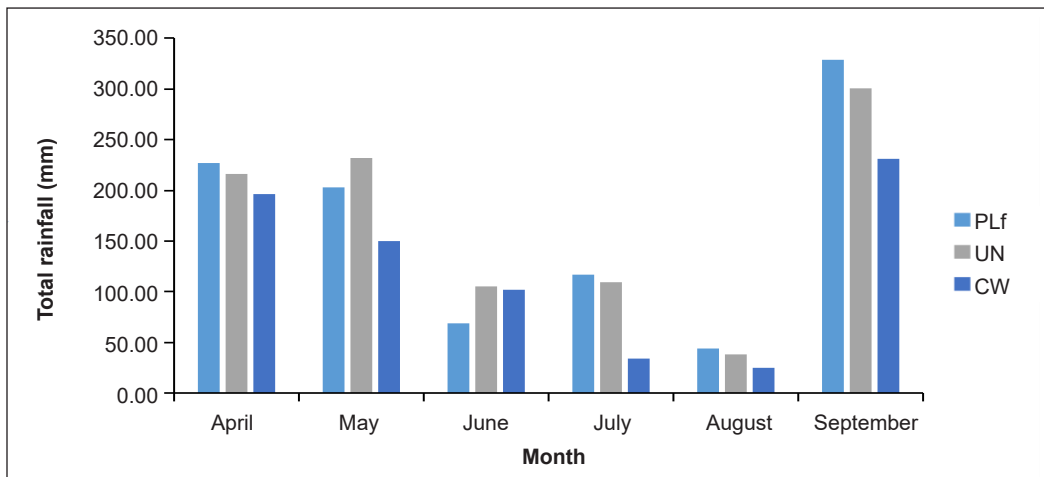

Figure 2. Rainfall pattern at three Putrajaya Lake and Putrajaya Wetlands Park sampling sites within the sampling months

Physico-chemical Analysis

In situ measurements were done for all physical parameters in accordance with the Standard Methods for the Examination of Water and Wastewater (APHA, 2005). Water salinity, conductivity, and temperature were assessed by a YSI multi-parameter probe (Model: 30 SCT Handheld meter). Dissolved oxygen was measured by YSI Dissolved Oxygen (Model 58), while pH was measured by Trans Instrument WalkLAB Microprocessor pH meter (HP9000). Irradiance was determined using a Digital Light Meter (Model: LD8903) and a Secchi disk measured water transparency. Nutrients (ammonia (NH_3), nitrate (NO_3^-), phosphate (PO_4^{3-}) and silica (Si)) were measured in the laboratory within the same sampling day by using Hach Multiparameter Portable Colorimeter (DR900, HACH). The chemical profiles of the water sample collected were assessed based on the instructions: nitrate

(Cadmium Reduction method: 8039), phosphate (Ascorbic acid method: 8048), ammonia (Salicylate method: 8155) and silica (Heteropoly Blue method: 8186) as described in the HACH handbook of Water Analysis (Hach, 2002). Readings for each parameter were recorded in triplicates. All equipment for *in situ* analysis was calibrated according to the manufacturer's instructions prior to sampling.

DNA Extraction

Triplicate samples of 500 mL water were filtered with Whatman polyethersulfone (PES) membrane filter paper with a pore size of 0.22 µm, 47 mm diameter, using a filtration pump. The filter papers containing the microbes were cut into small pieces using a sterile blade and stored in a -80°C freezer prior to DNA extraction. According to the manufacturer's suggested protocol, DNA was extracted using an EZNA Soil DNA Kit (Omega Bio-Tek, Inc.). The DNA purity analysis was carried out using a Nanodrop BioSpectrometer, and all samples showed a ratio value of A260/280 nm between the range of 1.8–2.0 (Kinetic, Eppendorf, Germany).

PCR Amplification and Illumina MiSeq Sequencing

The oligonucleotide primers used for PCR amplification targeting 16S V3-V4 regions in this study were (5'- CCT ACG GGN GGC WGC AG -3') and (5'-GAC TAC HVG GGT ATC TAA TCC -3') (Chen et al., 2019). The PCR was performed according to Illumina's 16S Metagenomic Sequencing Library protocol (Part # 15044223 Rev. B) (Illumina Inc., 2013). Each PCR reaction contained 25 µl of Master Mix (5 µl of Amplicon PCR Forward Primer (1 µM), 5 µl of Amplicon PCR Reverse Primer (1µM), 12.5 µl of 2× KAPA HiFi HotStart ReadyMix, and 2.5 µl of microbial genomic DNA (5 ng/µl in 10mM Tris pH 8.5). The PCR cycle conditions used in this study were as follows: 95°C for 3 minutes; 25 cycles of 95°C for 30 seconds, 55°C for 30 seconds, followed by 72°C for 30 seconds, with a final extension of 72°C for 5 minutes, and hold at 4°C. One µl of PCR products was run on 1.7% TAE agarose gel at 100V with 1500 bp DNA ladder for 65 mins. The gel was then viewed under a UV light gel documentation system (Enduro™ GDS). The PCR product size was approximately 430 bp. Amplicons were sent for Illumina Library Preparation and MiSeq sequencing at Apical Scientific Sdn. Bhd., Seri Kembangan.

Data Analysis

Apical Scientific Sdn. Bhd. provided all bioinformatics data analysis. The raw Illumina MiSeq sequence data were processed to remove sequence adaptors and low-quality reads using Bestus Bioinformatics Decontamination using Kmers (BBDuk) of the BBTools package (https://sourceforge.net/projects/bbmap/). Then, the paired-end reads (forward and reverse) were merged using USEARCH software (v11.0.667) (https://www.drive5.com/usearch/). Sequences shorter than 150 bp or longer than 600 bp were removed from

downstream processing to reduce the presence of unwanted adaptor and primer dimers (Gane, 2018) and increase the quality of reads, respectively (Tan et al., 2019). The remaining sequences were aligned against the 16S rRNA SILVA database (Release 132) and checked for chimeric errors using VSEARCH v2.6.2. In *de novo* OTU picking, all reads were clustered at 97% using UPARSE v11.0.667. Spurious OTUs with only one read (singleton) or two reads (doubleton) were removed from downstream processing to reduce the rate of sequencing errors and increase the accuracy of diversity metrics (Edgar, 2013; Flynn et al., 2015). For each OTU, a single representative sequence was chosen randomly, and Pynast software (https://www.ncbi.nlm.nih.gov/pubmed/19914921) was used to align and construct a phylogenetic tree against the SILVA 132 16S rRNA database. Benchmarking taxonomic assignment of OTUs was achieved using QIIME V1.9.1 against the Silva 16S rRNA database (release 132) (Kuczynski et al., 2012).

All statistical analyses were performed in R package V3.6. Alpha-diversity indices were calculated to assess species diversity in a sample using the UPARSE pipeline (Kuczynski et al., 2012), including the observed OTU, Chao1 and Shannon indices. The Vegan package performed the beta-diversity assessment to evaluate differences in taxonomic complexity among the samples using the phylogenetic weighted Unifrac distance (Lozupone et al., 2011) and ordination techniques using Principal coordinate analysis (PCoA) (Oksanen, 2008). Linear discriminant analysis effect size (LEfSe) was performed online (https://huttenhower.sph.harvard.edu/galaxy/) to determine the features of the bacterial group most likely to show differences between the two conditions or among the three sites based on the OTU lists of samples (Segata et al., 2011). For LEfSe analysis, an LDA score of >4 was used as a threshold. One-way Analysis of Similarity test (ANOSIM) was performed using the statistical software package PRIMER-E to test for significant differences between bacterial communities at three sites (Clarke & Gorley, 2006).

One-way analysis of variance (ANOVA) was performed in Statistical Package for Social Science (SPSS) version 25.0 to evaluate differences in physicochemical properties among the three sampling sites at a 0.05% significance level. A post hoc test (Tukey HSD) was used to compare means significantly different from each other at a probability of 5%. Principal component analysis (PCA) was used to examine the correlation among physicochemical parameters. Canonical correspondence analysis (CCA) was also performed to explore possible correlations between bacterial community and environmental variables using Microsoft Excel XLStat (Wan et al., 2017).

Data Availability

The raw 16S reads were deposited to NCBI Sequence Read Archive (SRA) under accession number PRJNA687158. All other data are provided in the main text or supplementary materials.

RESULTS

Physico-chemical Properties of the Samples

A comparison was made between the lake water input (UN) and the Central wetland (CW) due to the direct water flow from the inlet to CW for remediation. Water flowing from the wetland to the primary lake (PLf) was not significantly highlighted due to the distance (km) and various water inlets surrounding the lake, which might have an additional influence on the physicochemical parameters' readings.

Significant reduction of nutrients upon bioremediation at the wetland (CW) showed that the concentration of nitrate and silica was lower as the water passed through the Upper North (UN8) arm of the wetland (Figure 3). However, it was noted that there was no significant difference in physicochemical parameters between the dry and the wet (post-rainfall) event.
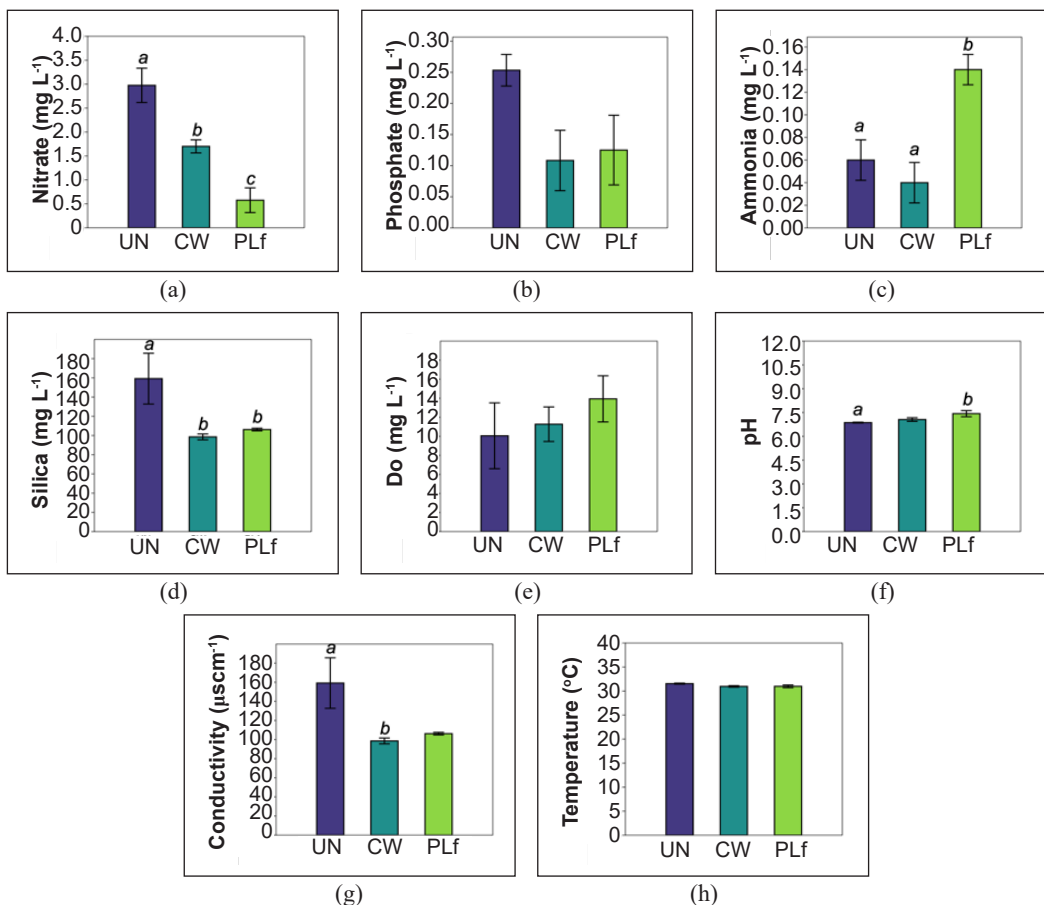

Figure 3. Mean physicochemical parameter measurements combined for May and July (A-G) in Upper North 8 (UN), Central Wetland (CW), Primary lakefront (PLf): (a) Nitrate; (b) Phosphate; (c) Ammonia; (d) Silica; (e) Dissolved Oxygen; (f) pH; (g) Conductivity; and (h) Temperature. Each bar represents the mean of triplicate data ± SE. Bars with different letters are significantly different (p<0.05)

Bacterial Community Sequencing, Data Rarefaction and Diversity Indices

For biological data, a comparison was made between May and July. In total, 3,474,032 raw reads were generated by Illumina MiSeq sequencing. After quality filtering, a total of 1,910,113 16S readings were obtained. Rarefaction analysis indicated that the number of OTUs reached the plateau phase (Supplementary Figure 1), suggesting sufficient sequence numbers to represent the actual bacterial community of specific sites and conditions. The value of diversity indices (Chao1, Shannon and observed OTU) of the bacterial community at the three sites was observed (Figure 4). The OTU number at the wetland in the wet condition (May) was 1248 on average, higher than the dry weather (July). The indices value of Chao1 and Shannon at the wetland in the dry event recorded the highest value, 1439 and 5.22 on average, respectively, while the inlet in the wet season recorded the highest index, 1233 and 4.88 on average.
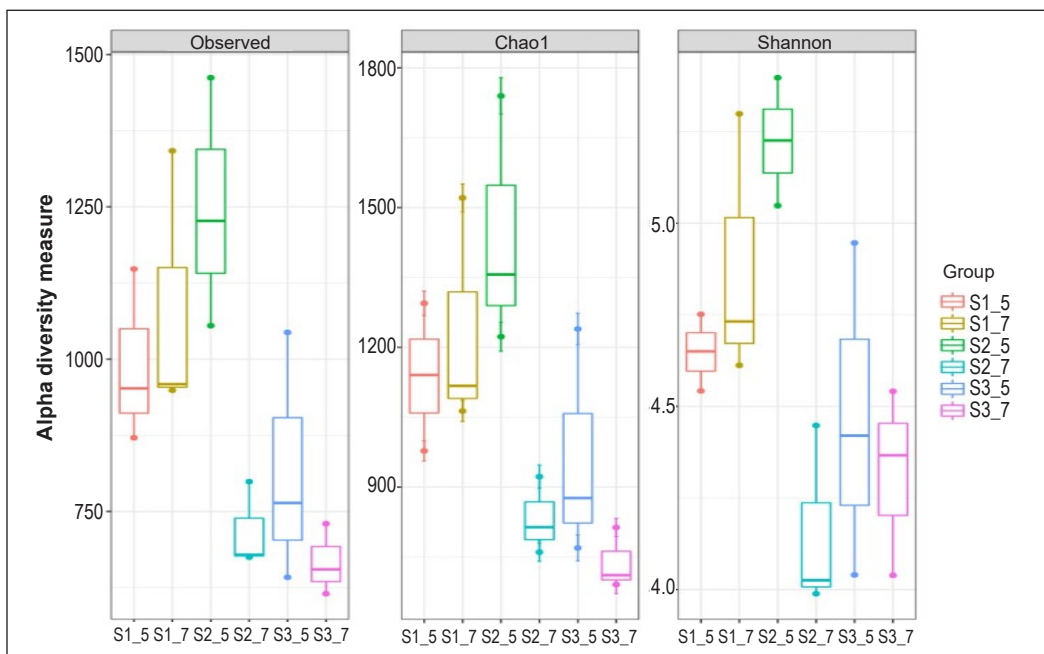

Figure 4. Comparison of diversity indices (OTU, Chao1 and Shannon). S1_5 (Inlet, wet), S1_7 (Inlet, dry), S2_5 (CW, wet), S2_7 (CW, dry), S3_5 (Lake, wet), S3_7 (Lake, dry)

Beta Diversity Differences Between Sites and Months

Beta diversity analysis was performed to analyse and compare the overall bacterial diversity within the community throughout the sampling period. Principal Coordinates Analysis (PCoA) plots following the Unique Fraction (UniFrac) method (Figure 5) showed the different compositions of the bacterial diversity within the samples. PC1 and PC2 explained 45.22% of the total microbial community structure.
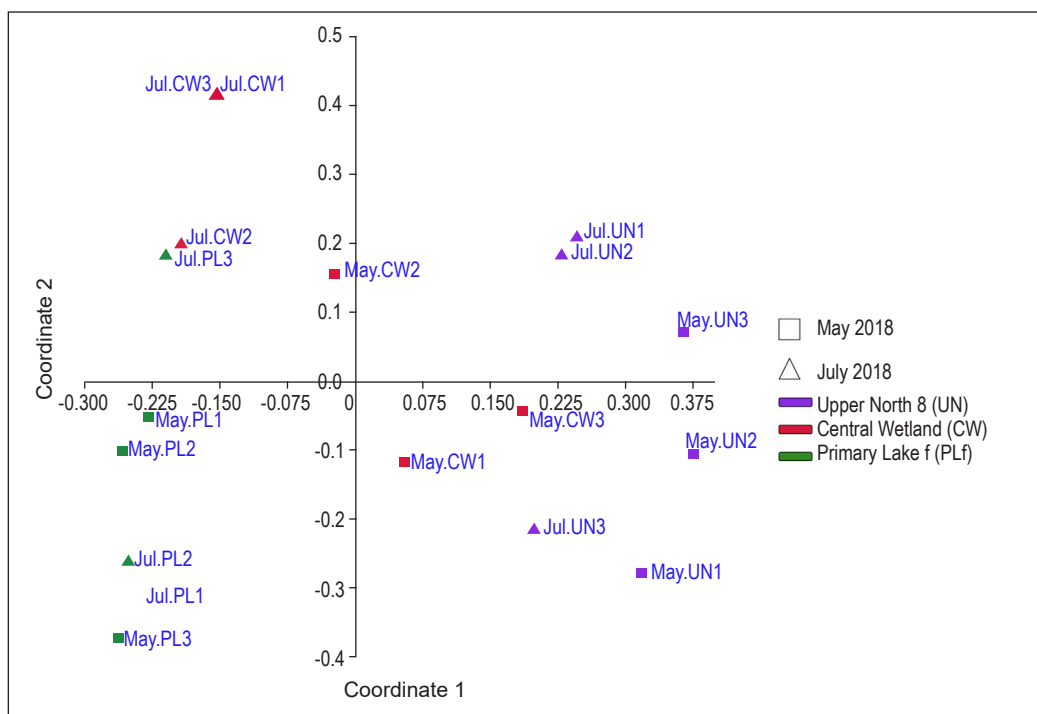
Figure 5. Principal coordinates analysis (PCoA) of bacterial communities. May: wet period, July: dry period. UN= Upper North 8, CW= Central Wetland, PL = Primary Lake f

PCoA identifies factors that differentiate the microbial communities. Clustering can be seen for PCoA ordination of samples of the primary lake (PL) communities as well as the upper North (UN) site. PCoA revealed seasonal community trends that were particularly apparent in PL and UN.

Weighted UniFrac distances were used to measure the dissimilarity coefficient between samples at different sampling sites based on the top ten taxa abundances. Unifrac distances showed that the diversity of bacterial OTUs was separated into two groups with different sampling times (Figure 6).

Bacterial Community Composition

The phylum and genus-level bacterial diversity distributions are shown in Figures 7 and 8, respectively. Forty-six major bacteria phyla were identified, including Proteobacteria (6.49%), Actinobacteria (2.95%), Bacteroidetes (2.86%), Verrucomicrobia (1.74%), Cyanobacteria (1.48%), Planctomycetes (1.34%), Firmicutes (0.52%), Chloroflexi (0.34%), Chlamydiae (0.14%) and Acidobacteria (0.09%). LEfSe analysis was performed to identify the most differentially abundant taxa between the two conditions by having a linear discriminant analysis (LDA) score of more than 4 (Supplementary Figures 2, 3 and 4). Proteobacteria were more abundant during the dry condition, including *Pseudomonas*
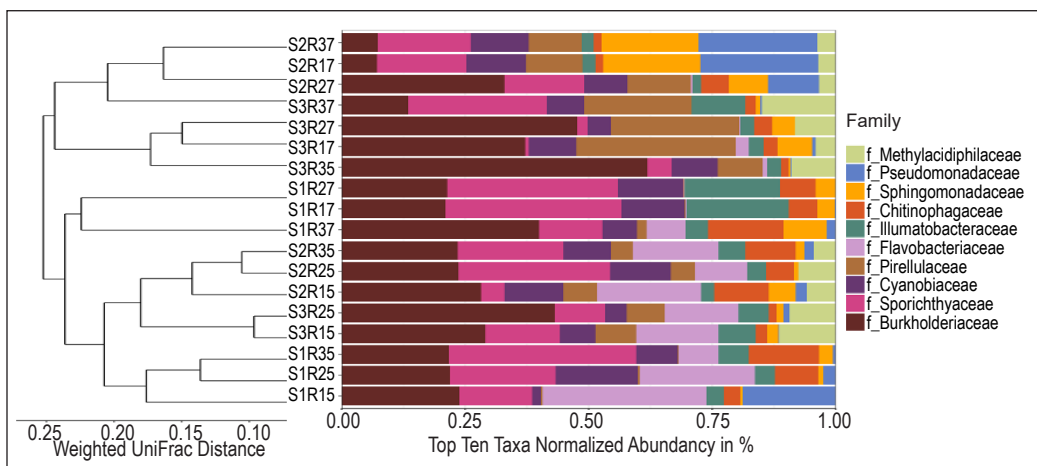
Figure 6. Weighted Unifrac dissimilarities of bacterial communities between samples at different sampling sites (S1: Inlet, S2: CW, S3: Lake, 5: Wet, 7: Dry)

sp., which were discovered in the dry condition at the wetland (0.92%) (Figures 7, 8 and Supplementary Table 2). In wet conditions, Bacteroidetes was the most common phylum (2.86%) at all three sites, with *Flavobacterium columnare* representing 2.54% of the total Bacteroidetes population. Planctomycetes were commonly found in dry conditions in the wetland, whereas the primary lake was represented by Pirellulales (Supplementary Figures 3 and 4). Phylum Firmicutes was discovered to be more common during the dry condition at both inlet and primary lake based on LEfSe analysis (Supplementary Figures 2 and 4). Upon the bioinformatics analysis against the database reference, we managed to identify 62 bacteria at the species level (Supplementary Table 1). The top 10 species and their presence in respective samples are represented in Table 2.

Table 2

Top 10 bacteria species discovered at Putrajaya Lake and Putrajaya Wetlands Park and presence in respective samples

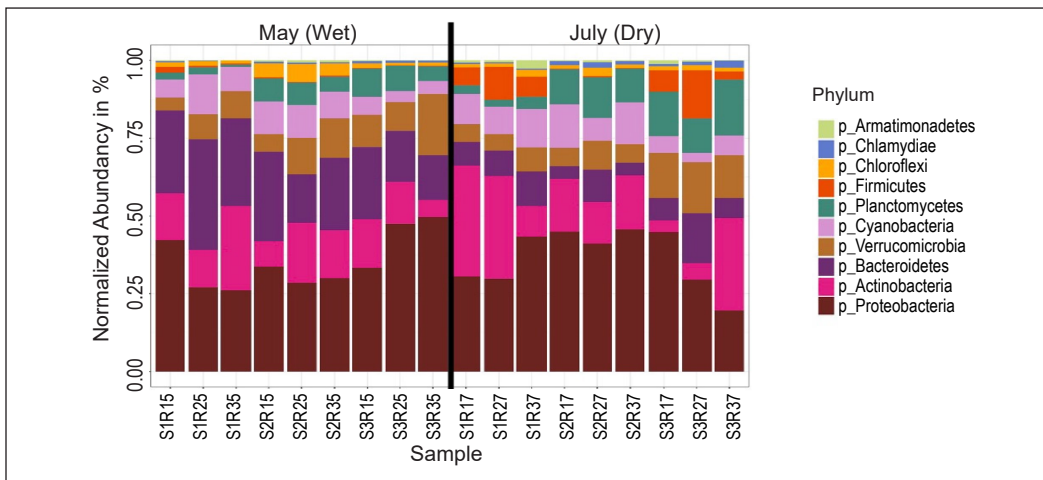| Bacteria	Sample					
	May 2018			July 2018		
	UN	CW	PL	UN	CW	PL
Flavobacterium columnare	/	/	/	/	/	X
Sphingomonas changbaiensis	X	/	/	/	/	/
Acidovorax delafieldii	/	/	/	/	/	/
Flavobacterium succinicans	/	/	/	/	/	X
Prosthecobacter debontii	/	/	/	/	/	/
Candidatus Aquiluna rubra	/	/	/	/	/	/
Sphingomonas wittichii	/	/	/	/	/	X
Roseateles depolymerans	/	/	/	/	/	/
Sphingomonas yabuuchiae	/	/	/	/	/	/
Pseudomonas nitroreducens	/	/	/	/	/	/

Figure 7. Phylum level distribution (%) of the bacterial community in Putrajaya Lake (S1: UN, S2: CW, S3: PLf, 5: Wet, 7: Dry)
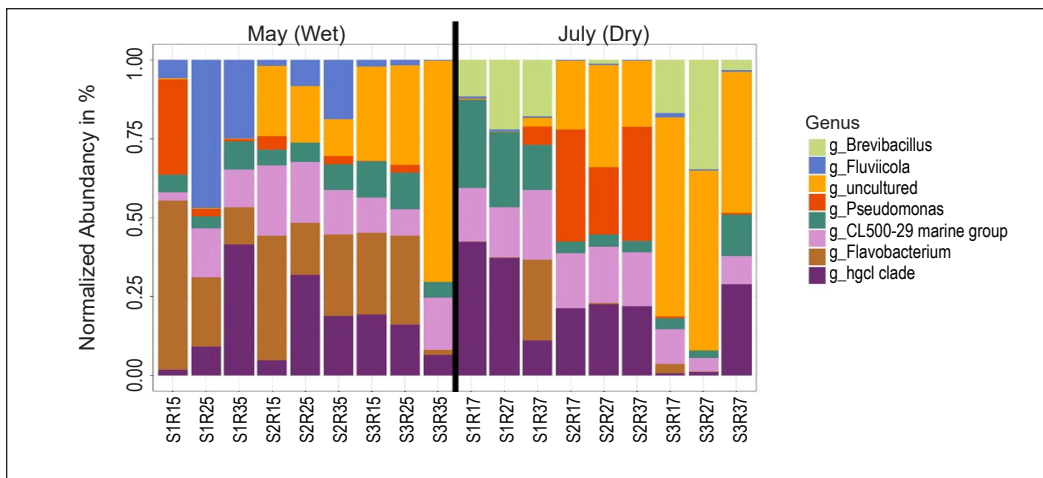

Figure 8. Genus level distribution (%) of the bacterial community in Putrajaya Lake (S1: UN, S2: CW, S3: PLf, 5: Wet, 7: Dry)

Relationships Between Bacterial Community Composition and Physico-chemical Properties

A PCA analysis examined correlations among the physicochemical parameters and sampling sites (Figure 9). The cumulative total variability of the data was 64.36%, explained by the first and second eigenvalues. PCA analysis was performed prior to the analysis of Canonical correspondence analysis (CCA) to extract precise parameters that affect bacterial distribution in Putrajaya Lake. PCA biplot shows both active variables (physicochemical parameters) as well as active observations (months and sampling sites).
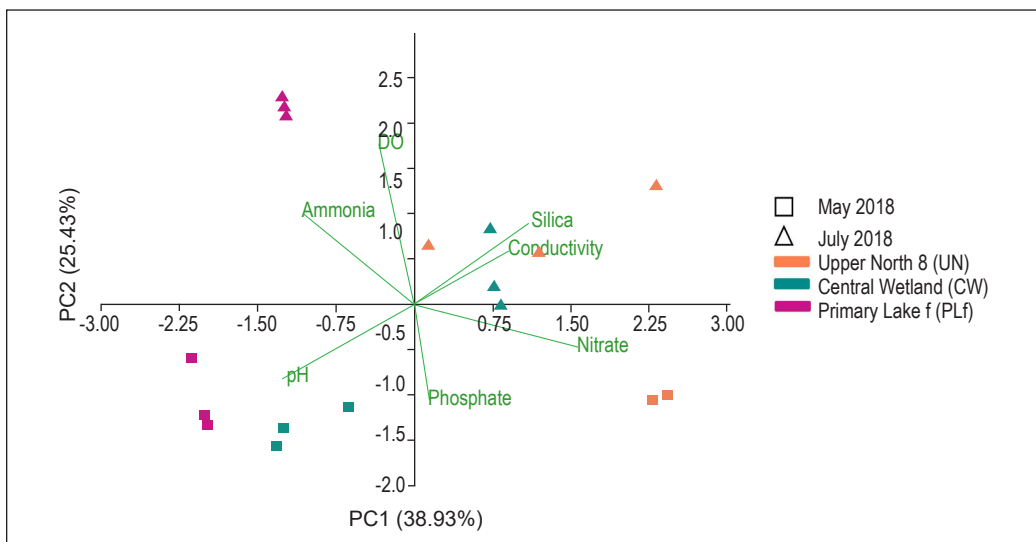
Figure 9. Principal component analysis (PCA) illustrates the correlation between variables and sites
Note. May: wet period. July: dry period

There was a positive correlation between nitrate and conductivity (0.52). A negative correlation was found between ammonia and nitrate (-0.72) and pH with nitrate (-0.64). pH was also negatively correlated with silica (-0.57). The post-rain fall condition (May samples) was more associated with pH, which can be observed in the central wetland and the primary lake.

ANOSIM analysis revealed significant differences in the bacterial communities among the three sites (R = 0.793, P = 0.001) (Figure 10). A one-way ANOSIM accepted the null hypothesis that there was a significant difference in total community structure at the level of individual sequences between sites. This finding may demonstrate a correlation between physicochemical parameters and bacterial composition.

The data in the box are the distances between and within groups R-value range (-1, 1). An R-value close to 0 represents no significant differences between and within groups, and an R-value close to 1 shows that within-group differences are greater than between-group differences. Boxes represent
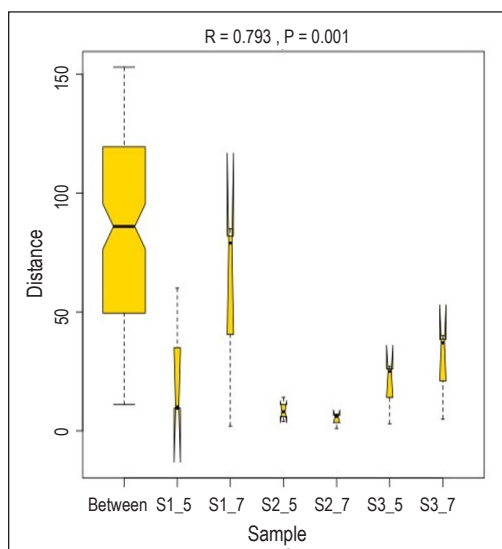

Figure 10. ANOSIM analysis results. Between represents the difference between groups (S1: UN, S2: CW, S3: PLf). Others are within groups 5: wet and 7: dry

the interquartile range between the first and third quartiles, while the line inside the box defines the median.

Pearson correlation revealed that Verrumicrobia and Firmicutes had a strong positive correlation with ammonia-nitrogen (r = 0.709). Actinobacteria and Cyanobacteria had a moderate positive correlation with nitrate with r value (r = 0.673) and (r = 0.647), respectively, while Bacteroidetes showed a positive association with temperature (r = 0.64) (Table 3).

CCA ordination plot (Figure 11) indicates the distribution of bacteria with the availability of physicochemical features at a site. This study's physicochemical parameters

Table 3

Pearson's correlation coefficient analysis of selected physicochemical variables and bacteria diversity from metagenomics data

Variables	pH	Nitrate	Phosphate	Ammonia	Temperature
Proteobacteria	0.44	-0.426	-0.199	-0.311	**-0.15**
Actinobacteria	-0.561	**0.673**	0.443	-0.205	**-0.07**
Bacteroidetes	0.144	0.266	0.424	-0.254	**0.64**
Verrumicrobia	**0.597**	**-0.781**	-0.327	**0.709**	**-0.32**
Cyanobacteria	**-0.612**	**0.647**	0.006	**-0.788**	**-0.12**
Planctomycetes	-0.368	0.446	-0.346	-0.194	**-0.83**
Firmicutes	-0.324	0.028	-0.152	**0.709**	**-0.41**
Chlorofelexi	-0.031	0.027	0.145	-0.162	**-0.01**
Acidobacteria	-0.113	0.325	**0.547**	-0.202	**0.26**
Chlamydiae	-0.174	-0.155	**-0.615**	-0.214	**-0.65**

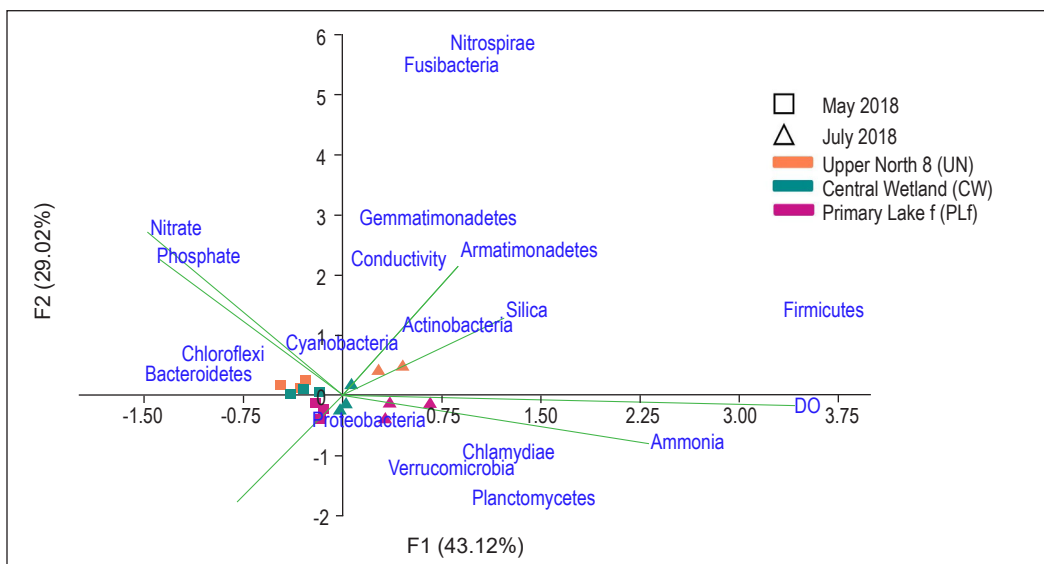*Note.* Values in bold significance level alpha=0.05

Figure 11. CCA ordination plot illustrating the correlation between variables and bacterial composition. May: wet period, July: dry period

and rainfall impacted the bacterial community composition. Overall, the first two axes explained 72.14% of the total variation. The diversity of bacteria among the samples was more clustered within the different sampling months than the sites. Samples from May were more driven by the nitrate and phosphate concentration, particularly associated with the community of Chloroflexi, Bacteroidetes and Cyanobacteria. Meanwhile, those from August were more influenced by the availability of silica and ammonia.

DISCUSSION

The advancement in sequencing methods through Next-Generation Sequencing (NGS) has been adopted as the main standard for bacterial identification, replacing the conventional culturing method and first-generation sequencing (Sanger sequencing). Implementing NGS for bacteria community profiling of freshwater lakes and water bodies in Malaysia has become increasingly popular as it provides much information for various research fields. A previous study on a tropical man-made lake in Malaysia, Lake Temenggor, discovered the influence of aquaculture activities on water quality by discovering the functional capacity of the bacteria profile through shotgun metagenomics (Lau et al., 2019). The same method was also carried out in a study for endophytic bacteria identification in Kenyir Lake, which may provide insightful information for antimicrobial properties among medicinal plants (Abidin et al., 2020). Bacterial communities are responsible for the biogeochemical cycling of nutrients in the ecosystem, particularly in the freshwater environment. This study was the first attempt to characterise the bacterial community based on molecular data in Putrajaya Lake within the dry season of May after rainfall events (wet) and in July (dry). The wet condition showed higher proportions of bacterial diversity than the dry condition, though they were not significantly different.

In eutrophic lakes, high phytoplankton concentration produces large amounts of organic compounds through photosynthesis and releases oxygen into the water (Viet et al., 2016). The dissolved oxygen is inversely proportional to the water temperature in the lake, which might explain the lower DO readings in the water inlet. However, bacteria that decompose phytoplankton organic compounds also use oxygen, causing net respiration (Kragh et al., 2020). Thus, the respiration of phytoplankton and bacteria which consume the oxygen can cause a decrease in the dissolved oxygen concentration in eutrophic lakes. Conductivity was also commonly used as a lake water quality indicator, which is proportional to the water temperature - a higher conductivity value indicates higher ionic concentrations. The pH of Putrajaya Lake is neutral to slightly alkaline. The alluvium soil, which can be found at the lake bed, consists of silty to sandy clay, which is low in bicarbonate, which may explain the overall neutral pH (Jaapar et al., 2002). The pH can fluctuate rapidly in aquatic ecosystems. High photosynthesis removes carbon dioxide and bicarbonate from the water, increasing carbonate concentrations and pH. The bacterial community may not adapt optimally to altered pH conditions during rapid pH changes, leading to impaired functions (Viet et al., 2016).

Nitrate and phosphate are among the chemical parameters used to assess water quality. The value of nitrate and silica decreased significantly from the water inlet near the Upper North 8 site (UN) to the central wetlands (CW), which may reflect the bioremediation impact of wetlands. In the central wetlands, plants associated with microbes (holobionts) could successfully reduce nitrate and silica concentrations. Holobionts can absorb nutrients such as nitrate, which cycles through several transformation processes, including fixation, ammonification, nitrification, and denitrification conducted by bacteria (Vymazal, 2007). Holobionts are chosen based on soil type and water level in the Putrajaya wetland area. *Hanguana malayana* (Bakong) and *Phragmites karka* (Rumput gedabong) are the most suitable plants found to be grown in wetland areas from 1998 until now (Mohamad, 2012). Silica is an essential element for the growth of diatoms. Aquatic macrophytes like *Phragmites australis* (reed) can uptake, translocate and accumulate silica in wetland areas (Struyf et al., 2004). This species can temporarily store silica, making it unavailable for diatoms. Ammonia-nitrogen can be used by phytoplankton or oxidised by bacteria into nitrate or nitrite for their growth and metabolism (Glibert et al., 2016). The higher concentrations of ammonia-nitrogen were influenced by the possible discharge of domestic sewage, agricultural activities and animal waste (Huang et al., 2015).

Importance and Application of Bacterial Species Found in Putrajaya Lake and Putrajaya Wetlands Park

Proteobacteria were more common in the dry condition than the wet condition based on LEfSe analysis. Soil bacteria such as Proteobacteria are important drivers of energy flow and nutrient cycling from the physical environment into living organisms in lake ecosystems (Wan et al., 2017). High abundances of Proteobacterial groups have previously been related to the higher availability of carbon (Fierer et al., 2007). Photosynthesis by aquatic plants and phytoplankton is a major carbon source in lakes. Aquatic ecosystems also get carbon sources through the runoff from the land nearby. In this study, the availability of carbon in the neutral pH range to slightly more alkaline might have contributed to the change in the relative abundance of Proteobacteria, which include *Pseudomonas* sp., which might have originated from non-point sources such as agricultural and domestic. *Pseudomonas* sp. is able to degrade organic compounds into simpler molecules in a variety of ecological conditions (Batrich et al., 2019). The availability of these microbes in the constructed wetland may be involved in nutrient cycling, filtering and removing pollutants coming into the lake.

The eutrophication of the lake enriched with sediments is preferred by nutrient-cycling bacteria such as Bacteroidetes, Firmicutes, Alphaproteobacteria and Gammaproteobacteria (Huang et al., 2017). *Sphingomonas sp.*, an example of the Alphaproteobacteria, were abundant and capable of surviving in various environments. *Sphingomonas* are known for degrading hydrocarbons, which can aid in bioremediation. The *Sphingomonas* species

discovered within Putrajaya Lake and Putrajaya Wetlands Park have been previously reported to be capable of degrading hydrocarbons to some degree. It allows for the degradation of pollutants such as petrol spills and other oil-based spills that are potentially harmful to the ecosystem. *Sphingomonas changbaiensis*, as an example, is often added as parts of bacterial consortia used together with biosurfactants such as alkyl polyglycoside, which enhances the rate of degradation of total petroleum hydrocarbons (Li et al., 2020). *Sphingomonas wittichii*, specifically strain RW1, is a highly-interest bacterium due to its ability to degrade dioxin-based pollutants such as polychlorinated (Colquhoun et al., 2012).

The highest operational taxonomic unit (OTU) number assigned at the species level in this study was a Gram-negative Bacteroidetes under the class Flavobacteria. *Flavobacterium* sp. is a ubiquitous fish pathogen genus, notably within freshwater habitats. It can cause infections in different parts of fish (Loch & Faisal, 2015). The *Flavobacterium columnare* is responsible for a prevalent fish disease known as columnaris or cotton mouth disease, which has been reported to infect fishes in wild habitats and cultures. Previous studies reported infection by this bacterial pathogen at a global scale among various fish species (Davis, 2011), such as trout (Singh et al., 2021), red tilapia (Ponpukdee et al., 2021), Asian seabass (Chokmangmeepisarn et al., 2021) and catfish (Lange et al., 2021). This species has also been reported to interact with the parasitic protozoan *Chilodonella hexasticha* to cause fish death. This ciliate is correlated with the abundant presence of *F. columnare* (Gomes et al., 2019).

Another species, *F. succinicans,* is associated with gill disease among rainbow trout (Good et al., 2015). In Malaysia, the presence of *Flavobacterium* sp. is widely spread in freshwater habitats to the extent that some may have developed antibiotic resistance (Hassan et al., 2020). A study among the ornamental fish industry in Malaysia has also found the presence of this genus infecting fishes sold in local retail pet shops (Anjur et al., 2021). The treatment for *Flavobacterium* infection is currently by administering vaccines or antibiotics to the culture (Lange et al., 2019). Though these species are not a concern for human health, more strains have become resistant to the treatments (Elgendy et al., 2022), causing a massive economic loss in the aquaculture sector (LaFrentz et al., 2018).

The distribution of the *Flavobacterium* population within the aquatic habitat is reportedly associated with heterotrophic activity within the environment, favouring the abundance of available resources (Newton et al., 2011). The ability to survive and quickly adapt to high nutrient concentrations might explain the high OTUs assigned to *Flavobacterium* reported in this study across all sampling sites.

Ecohydrology management services by the Putrajaya Corporation (PjC) encompass maintenance activities, monitoring exercises and biodiversity assessment. This research will contribute towards the monitoring exercises of the physical and chemical properties of water. The report on the diversity of bacterial communities from the water input site (Upper North: UN8), remediation site (Central Wetland: CW), as well as the main water

sports activities centre (Primary Lake f: PLf), provide specific snapshots of spatiotemporal bacterial community data which will contribute towards Putrajaya's biodiversity and ecological assessment. Water flowing from the Upper North to the primary lake area showed a marked decrease in certain nutrient concentrations, providing evidence of the effectiveness of the engineered wetlands of Putrajaya.

Implications for Remediation Strategies

As a constructed lake that is continuously exposed to various anthropogenic pollutants coming from untreated water of the two main rivers, there is a need to develop efficient strategies to ensure that the Putrajaya Lake and its surrounding water inputs are of good quality, on par with the level set by the National Water Quality Standard for Malaysia (NWQSM) and maintained at pristine condition. One of the major issues in Putrajaya Lake is the accidental release of nutrients in runoffs, such as landscape fertilisers and pesticides, which cause the water clarity to decrease and result in consistent slight eutrophic conditions (Hakim et al., 2016). The lake also serves as a major tourism site for water-based activities. Hence, a strict monitoring program is essential to sustain and maintain the water quality. Pollutants come from point and non-point sources, inside and outside the boundaries of Putrajaya Lake; thus, the multiple stakeholders' management is essential, in addition to the local government. Therefore, a holistic management approach needs to be emphasised by all stakeholders.

CONCLUSION

The use of the 16S rRNA gene as the genetic marker was successful in identifying bacteria species. Overall, 49 phyla, 147 classes, 284 orders, 471 families, 778 genera and 62 species of bacteria were identified. Proteobacteria was the most dominant bacteria class found in all samples, while *Flavobacterium columnare* of Bacteroidetes was the most common bacteria to be identified at the species level. Verrumicrobia and Firmicutes showed a strong positive correlation with ammonia-nitrogen ($r = 0.709$). Actinobacteria and Cyanobacteria had a moderate positive correlation with nitrate with r value ($r = 0.673$) and ($r = 0.647$), respectively. This report has been the first to employ the 16S rRNA gene amplicons to evaluate the bacterial community assemblages of Putrajaya Lake, Malaysia. We have found that the presence of the engineered wetland plantings was influential in nutrient removal, as shown by the significant reduction in nitrate and silica concentration, which might have impacted the diversity of bacteria found within the different sampling sites. For future research, other wetland arms, such as the Upper Bisa (UB), Upper West (UW) and Upper East (UE), will be considered for sampling sites. Engineered wetlands have benefited the natural bioremediation system of lake water inlets and could be emulated in other cities for environmentally friendly water management.

Nurul Najwa Farihah Mat Lazim, Afiqah Mohamed, Zana Ruhaizat Zana Rudin, Fatimah Md Yusoff, Ikhsan Natrah and Shahrizim Zulkifly

ACKNOWLEDGEMENTS

REFERENCES

Abidin, Z. A. Z., Zolkefli, N., Shahari, R., & Chowdhury, A. J. K. (2020). Endophytic bacteria of *Anisophyllea disticha* (Raja berangkat) from tropical lake environment in Malaysia. *Desalination and Water Treatment*, *188*, 282-287. https://doi.org/10.5004/dwt.2020.25274

Al-Sulaiman, S. A. A. W. (2012). Chemical safety of drinking-water: Assessing priorities for risk management. *International Journal of Environmental Studies, 69*(6), 1001. https://doi.org/10.1080/00207233.2011.565947

Anjur, N., Sabran, S. F., Daud, H. M., & Othman, N. Z. (2021). An update on the ornamental fish industry in Malaysia: *Aeromonas hydrophila*-associated disease and its treatment control. *Veterinary World*, *14*(5), 1143-1152. https://doi.org/10.14202/vetworld.2021.1143-1152

APHA. (2005). *Standard Methods for the Examination of Water and Wastewater* (21st ed.). Washington DC. American Public Health Association.

Asmat, A., Hazali, N. A., Nor, A. N. M., & Zuhan, F. K. (2018). Seasonal-spatial of Putrajaya Lake Water Quality Parameter (WQP) concentration using geographic information system (GIS). *International Journal of Engineering & Technology*, *7*(3), 176-181. https://doi.org/10.14419/ijet.v7i3.11.15956

Batrich, M., Maskeri, L., Schubert, R., Ho, B., Kohout, M., Abdeljaber, M., Abuhasna, A., Kholoki, M., Psihogios, P., Razzaq, T., Sawhney, S., Siddiqui, S., Xoubi, E., Cooper, A., Hatzopoulos, T., & Putonti, C. (2019). Pseudomonas diversity within urban freshwaters. *Frontiers in Microbiology*, *10*, Article 195. https://doi.org/10.3389/fmicb.2019.00195

Briones, A., & Raskin, L. (2003). Diversity and dynamics of microbial communities in engineered environments and their implications for process stability. *Current Opinion in Biotechnology, 14*(3), 270-276. https://doi.org/10.1016/S0958-1669(03)00065-X

Chen, Z., Hui, P. C., Hui, M., Yeoh, Y. K., Wong, P. Y., Chan, M. C. W., Wong, M. C. S., Ng, S. C., Chan, F. K. L., & Chan, P. K. S. (2019). Impact of preservation method and 16S rRNA hypervariable region on gut microbiota profiling. *MSystems, 4*(1), 1-15. https://doi.org/10.1128/msystems.00271-18

Chokmangmeepisarn, P., Thangsunan, P., Kayansamruaj, P., & Rodkhum, C. (2021). Resistome characterization of *Flavobacterium columnare* isolated from freshwater cultured Asian sea bass (*Lates calcarifer*) revealed diversity of quinolone resistance associated genes. *Aquaculture*, *544*, Article 737149. https://doi.org/10.1016/j.aquaculture.2021.737149

Clarke, K. R., & Gorley, R. N. (2006). *Primer v. 6: User manual/Tutorial.* Scientific Research Publishing Inc. https://www.scirp.org/(S(vtj3fa45qm1ean45vvffcz55))/reference/ReferencesPapers.aspx?ReferenceID=1960070

Colquhoun, D. R., Hartmann, E. M., & Halden, R. U. (2012). Proteomic profiling of the dioxin-degrading bacterium *Sphingomonas wittichii* RW1. *Journal of Biomedicine and Biotechnology*, *2012*, Article 408690. https://doi.org/10.1155/2012/408690

Cotner, J. B., & Biddanda, B. A. (2002). Small players, large role: Microbial influence on biogeochemical processes in pelagic aquatic ecosystems. *Ecosystems, 5*(2), 105-121. https://doi.org/10.1007/s10021-001-0059-3

Davis, H. S. (2011). *A New Bacterial Disease of Fresh-Water Fishes*. US Government Printing Office. https://doi.org/10.5962/bhl.title.49773

Edgar, R. C. (2013). UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, *10*, 996-998. https://doi.org/10.1038/nmeth.2604

Elgendy, M. Y., Abdelsalam, M., Mohamed, S. A., & Ali, S. E. (2022). Molecular characterization, virulence profiling, antibiotic susceptibility, and scanning electron microscopy of *Flavobacterium columnare* isolates retrieved from Nile tilapia (*Oreochromis niloticus*). *Aquaculture International*, *30*, 845-862. https://doi.org/10.1007/s10499-021-00819-x

Fadrosh, D. W., Ma, B., Gajer, P., Sengamalay, N., Ott, S., Brotman, R. M., & Ravel, J. (2014). An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome*, *2*(6), 1-7. https://doi.org/10.1186/2049-2618-2-6

Fierer, N., Bradford, M. A., & Jackson, R. B. (2007). Toward an ecological classification of soil bacteria. *Ecology*, *88*(6), 1354-1364. https://doi.org/10.1890/05-1839

Flynn, J. M., Brown, E. A., Chain, F. J. J., Macisaac, H. J., & Cristescu, M. E. (2015). Toward accurate molecular identification of species in complex environmental samples: Testing the performance of sequence filtering and clustering methods. *Ecology and Evolution*, *5*(11), 2252-2266. https://doi.org/10.1002/ece3.1497

Gane, A. (2018). *Size selection brings better data to NGS workflows.* Cytiva. https://www.cytivalifesciences.com/en/us/news-center/better-data-via-size-selection-10001

Glibert, P. M., Wilkerson, F. P., Dugdale, R. C., Raven, J. A., Dupont, C. L., Leavitt, P. R., Parker, A. E., Burkholder, J. M., & Kana, T. M. (2016). Pluses and minuses of ammonium and nitrate uptake and assimilation by phytoplankton and implications for productivity and community composition, with emphasis on nitrogen-enriched conditions. *Limnology and Oceanography*, *61*(1), 165-197. https://doi.org/10.1002/lno.10203

Gomes, G. B., Hutson, K. S., Domingos, J. A., Villamil, S. I., Huerlimann, R., Miller, T. L., & Jerry, D. R. (2019). Parasitic protozoan interactions with bacterial microbiome in a tropical fish farm. *Aquaculture*, *502*, 196-201. https://doi.org/10.1016/j.aquaculture.2018.12.037

Good, C., Davidson, J., Wiens, G. D., Welch, T. J., & Summerfelt, S. (2015). *Flavobacterium branchiophilum* and *F. succinicans* associated with bacterial gill disease in rainbow trout *Oncorhynchus mykiss* (Walbaum) in water recirculation aquaculture systems. *Journal of Fish Diseases*, *38*(4), 409-413. https://doi.org/10.1111/jfd.12249

Gupta, V., Ratha, S. K., Sood, A., Chaudhary, V., & Prasanna, R. (2013). New insights into the biodiversity and applications of cyanobacteria (blue-green algae) - Prospects and challenges. *Algal Research, 2*(2), 79-97. https://doi.org/10.1016/j.algal.2013.01.006

Hach, C. (2002). *Water Analysis Handbook*. Loveland Publishers & Pioneers.

Hakim, M., Heryansyah, A., & Ismail, T. (2016, June 7-8). Sustainability in water resources management in Putrajaya wetland: Impact assessment on new emerging pollutants. In *1st Proceeding of Civil Engineering* (pp. 301-317). Universiti Teknologi Malaysia, Malaysia.

Hassan, N. L., Kam, K. Y., & Zain, N. A. M. (2020). Isolation of antibiotic resistant bacteria from rivers in Kelantan, Malaysia. *International Journal of Life Sciences and Biotechnology*, *3*(2), 291-307. https://doi.org/10.38001/ijlsb.712542

Huang, W., Chen, X., Jiang, X., & Zheng, B. (2017). Characterization of sediment bacterial communities in plain lakes with different trophic statuses. *MicrobiologyOpen*, *6*(5), 1-14. https://doi.org/10.1002/mbo3.503

Huang, Y. F., Ang, S. Y., Lee, K. M., & Lee, T. S. (2015). Quality of water resources in Malaysia. In *Research and Practices in Water Quality* (pp. 65-94). InTech. https://doi.org/10.5772/58969

Iliev, I., Yahubyan, G., Marhova, M., Apostolova, E., Gozmanova, M., Gecheva, G., Kostadinova, S., Ivanova, A., & Baev, V. (2017). Metagenomic profiling of the microbial freshwater communities in two Bulgarian reservoirs. *Journal of Basic Microbiology*, *57*(8), 669-679. https://doi.org/10.1002/jobm.201700137

Jaapar, A. R., Ahba, N. A., & Hussin, A. (2002, May 26). The occurrence and classification of hard rock body in Putrajaya and its implication to construction activities. *Bulletin of the Geological Society of Malaysia, 45*, 123-128. https://doi.org/10.7186/bgsm45200218

Jamal, A., Yusoff, F. M., Banerjee, S., & Shariff, M. (2014). Littoral and limnetic phytoplankton distribution and biodiversity in a tropical man-made lake, Malaysia. *Advanced Studies in Biology, 6*(4), 149-168. http://dx.doi.org/10.12988/asb.2014.4631

Jo, J. H., Kennedy, E. A., & Kong, H. H. (2016). Research techniques made simple: Bacterial 16S ribosomal RNA gene sequencing in cutaneous research. *Journal of Investigative Dermatology, 136*(3), e23-e27. https://doi.org/10.1016/j.jid.2016.01.005

Kragh, T., Martinsen, K. T., Kristensen, E., & Sand-Jensen, K. (2020). From drought to flood: Sudden carbon inflow causes whole-lake anoxia and massive fish kill in a large shallow lake. *Science of the Total Environment, 739*, Article 140072. https://doi.org/10.1016/j.scitotenv.2020.140072

Kuczynski, J., Stombaugh, J., Walters, W. A., González, A., Caporaso, J. G., & Knight, R. (2012). Using QIIME to analyze 16s rRNA gene sequences from microbial communities. *Current Protocols in Microbiology, 27*(1), 1E.5.1-1E.5.20. https://doi.org/10.1002/9780471729259.mc01e05s27

LaFrentz, B. R., García, J. C., Waldbieser, G. C., Evenhuis, J. P., Loch, T. P., Liles, M. R., Wong, F. S., & Chang, S. F. (2018). Identification of four distinct phylogenetic groups in Flavobacterium columnare with fish host associations. *Frontiers in Microbiology*, *9*, Article 452. https://doi.org/10.3389/fmicb.2018.00452

Lange, M. D., Abernathy, J., & Farmer, B. D. (2019). Evaluation of a recombinant flavobacterium columnare DnaK protein vaccine as a means of protection against columnaris disease in channel catfish (*Ictalurus punctatus*). *Frontiers in Immunology*, *10*, Article 1175. https://doi.org/10.3389/fimmu.2019.01175

Lange, M. D., Abernathy, J., Farmer, B. D., & Beck, B. H. (2021). Use of an immersion adjuvant with a Flavobacterium columnare recombinant protein vaccine in channel catfish. *Fish and Shellfish Immunology*, *117*, 136-139. https://doi.org/10.1016/j.fsi.2021.07.025

Lau, N. S., Zarkasi, K. Z., Sah, A. S. R. M., & Shu-Chien, A. C. (2019). Diversity and coding potential of the microbiota in the photic and aphotic zones of tropical man-made lake with intensive aquaculture activities: A case study on Temengor Lake, Malaysia. *Microbial Ecology*, *78*, 20-32. https://doi.org/10.1007/s00248-018-1283-0

Li, Q., Huang, Y., Wen, D., Fu, R., & Feng, L. (2020). Application of alkyl polyglycosides for enhanced bioremediation of petroleum hydrocarbon-contaminated soil using *Sphingomonas changbaiensis* and *Pseudomonas stutzeri*. *Science of the Total Environment*, *719*, Article 137456. https://doi.org/10.1016/j.scitotenv.2020.137456

Liles, M. R., Manske, B. F., Bintrim, S. B., Handelsman, J., & Goodman, R. M. (2003). A census of rRNA genes and linked genomic sequences within a soil metagenomic library. *Applied and Environmental Microbiology*, *69*(5), 2684-2691. https://doi.org/10.1128/AEM.69.5.2684-2691.2003

Loch, T. P., & Faisal, M. (2015). Emerging flavobacterial infections in fish: A review. *Journal of Advanced Research, 6*(3), 283-300. https://doi.org/10.1016/j.jare.2014.10.009

Lozupone, C., Lladser, M. E., Knights, D., Stombaugh, J., & Knight, R. (2011). UniFrac: An effective distance metric for microbial community comparison. *The ISME Journal, 5*, 169-172. https://doi.org/10.1038/ismej.2010.133

Majizat, A., Ahmad, B., Noordin, N., & Sharip, Z. (2016). Adopting an integrated catchment-based management approach for Putrajaya Lake and Wetlands. *International Journal of Water, 10*(4), 343-358. https://doi.org/10.1504/IJW.2016.079701

Mohamad, N. H. N. (2012). Ecological approaches in designing neighbourhood green spaces as urban wildlife habitat in the Klang valley, Peninsular Malaysia. *International Journal of Applied Science and Technology, 2*(3), 192-213.

Mueller-Spitz, S. R., Goetz, G. W., & McLellan, S. L. (2009). Temporal and spatial variability in nearshore bacterioplankton communities of Lake Michigan. *FEMS Microbiology Ecology*, *67*(3), 511-522. https://doi.org/10.1111/j.1574-6941.2008.00639.x

Najah, A., Teo, F. Y., Chow, M. F., Huang, Y. F., Latif, S. D., Abdullah, S., Ismail, M., & El-Shafie, A. (2021). Surface water quality status and prediction during movement control operation order under COVID-19 pandemic: Case studies in Malaysia. *International Journal of Environmental Science and Technology*, *18*, 1009-1018. https://doi.org/10.1007/s13762-021-03139-y

Newton, R. J., Jones, S. E., Eiler, A., McMahon, K. D., & Bertilsson, S. (2011). A guide to the natural history of freshwater lake bacteria. *Microbiology and Molecular Biology Reviews*, *75*(1), 14-49. https://doi.org/10.1128/mmbr.00028-10

Oksanen, J. (2008). Vegan: An Introduction to Ordination. https://cran.r-project.org/web/packages/vegan/vignettes/intro-vegan.pdf

Pernthaler, J., & Amann, R. (2005). Fate of heterotrophic microbes in pelagic habitats: Focus on populations. *Microbiology and Molecular Biology Reviews*, *69*(3), 440-461. https://doi.org/10.1128/mmbr.69.3.440-461.2005

Ponpukdee, N., Wangman, P., Rodkhum, C., Pengsuk, C., Chaivisuthangkura, P., Sithigorngul, P., & Longyant, S. (2021). Detection and identification of a fish pathogen *Flavobacterium columnare*

using specific monoclonal antibodies. *Aquaculture*, *545*, Article 737231. https://doi.org/10.1016/j.aquaculture.2021.737231

Sabkie, N., Zulkifly, S., Lazim, N., Rosli, N., & Yong, C. (2020). Physico-chemical influence on the diversity of phytoplankton at Putrajaya Lake and Wetlands, Putrajaya, Malaysia. *Malayan Nature Journal*, *72*(2), 153-163.

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., & Huttenhower, C. (2011). Metagenomic biomarker discovery and explanation. *Genome Biology*, *12*, Article R60. https://doi.org/10.1186/gb-2011-12-6-r60

Sharip, Z., Saman, J. M., Noordin, N., Majizat, A., Suratman, S., & Shaaban, A. J. (2016). Assessing the spatial water quality dynamics in Putrajaya Lake: A modelling approach. *Modeling Earth Systems and Environment*, *2*, Article 46. https://doi.org/10.1007/s40808-016-0104-z

Singh, S., Mallik, S. K., Kala, K., Shahi, N., Pathak, R., Giri, A. K., Chandra, S., Pant, K., & Patiyal, R. S. (2021). Characterization of *Flavobacterium columnare* from farmed infected rainbow trout, *Oncorhynchus mykiss* (Walbaum, 1792) of Central Indian Himalayan region, India. *Aquaculture*, *544*, Article 737118. https://doi.org/10.1016/j.aquaculture.2021.737118

Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., Arrieta, J. M., & Herndl, G. J. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere." *Proceedings of the National Academy of Sciences of the United States of America*, *103*(32), 12115-12120. https://doi.org/10.1073/pnas.0605127103

Staley, C., Unno, T., Gould, T. J., Jarvis, B., Phillips, J., Cotner, J. B., & Sadowsky, M. J. (2013). Application of Illumina next-generation sequencing to characterize the bacterial community of the Upper Mississippi River. *Journal of Applied Microbiology*, *115*(5), 1147-1158. https://doi.org/10.1111/jam.12323

Struyf, E., van Damme, S., & Meire, P. (2004). Possible effects of climate change on estuarine nutrient fluxes: A case study in the highly nutrified Schelde estuary (Belgium, the Netherlands). *Estuarine, Coastal and Shelf Science*, *60*(4), 649-661. https://doi.org/10.1016/j.ecss.2004.03.004

Suratman, S., Hussein, A. N. A. R., Tahir, N. M., Latif, M. T., Mostapa, R., & Weston, K. (2016). Seasonal and spatial variability of selected surface water quality parameters in Setiu Wetland, Terengganu, Malaysia. *Sains Malaysiana*, *45*(4), 551-558.

Tan, G., Opitz, L., Schlapbach, R., & Rehrauer, H. (2019). Long fragments achieve lower base quality in Illumina paired-end sequencing. *Scientific Reports*, *9*, Article 2856. https://doi.org/10.1038/s41598-019-39076-7

Viet, N. D., Bac, N. A., & Huong, H. T. T. (2016). Dissolved oxygen as an indicator for eutrophication in freshwater lakes. In *Proceedings of International Conference on Environmental Engineering and Management for Sustainable Development* (pp. 1-6). ResearchGate.

Vymazal, J. (2007). Removal of nutrients in various types of constructed wetlands. *Science of the Total Environment*, *380*(1-3), 48-65. https://doi.org/10.1016/j.scitotenv.2006.09.014

Wan, Y., Ruan, X., Zhang, Y., & Li, R. (2017). Illumina sequencing-based analysis of sediment bacteria community in different trophic status freshwater lakes. *MicrobiologyOpen*, *6*(4), Article e00450. https://doi.org/10.1002/mbo3.450

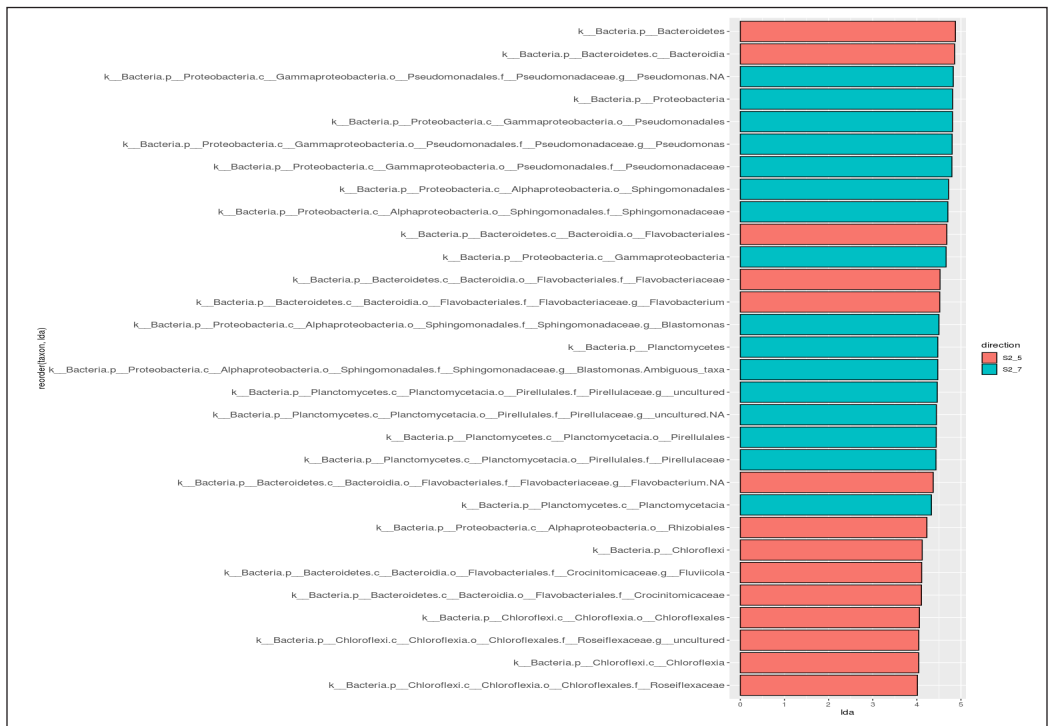# APPENDICES
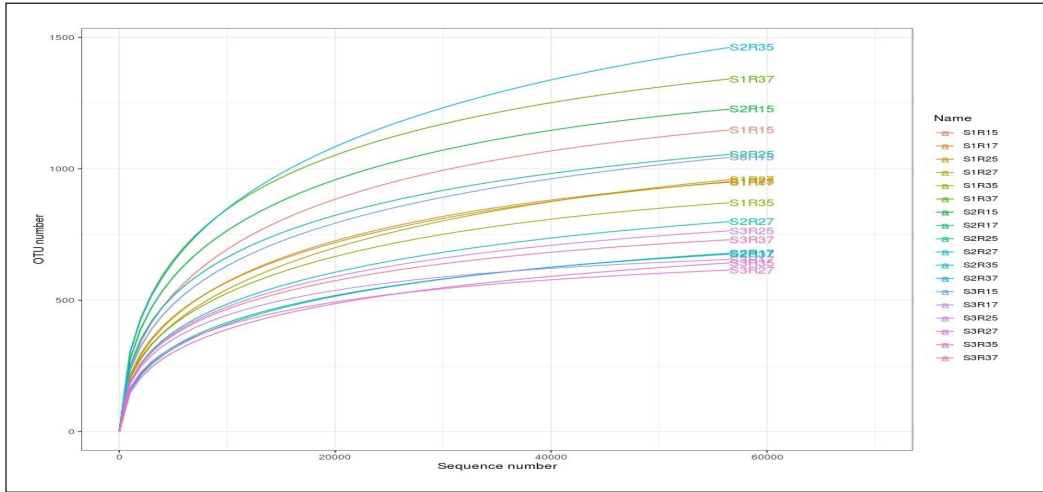
Supplementary Figure 1. Rarefaction curve. Wet season (May), Site 1 (S1R15, S1R25, S1R35; Site 2 (S2R15, S2R25, S2R35) and Site 3 (S3R15, S3R25, S3R35); dry season (July), Site 1 (S1R17, S1R27, S1R37; Site 2 (S2R17, S2R27, S2R37) and Site 3 (S3R17, S3R27, S3R37)
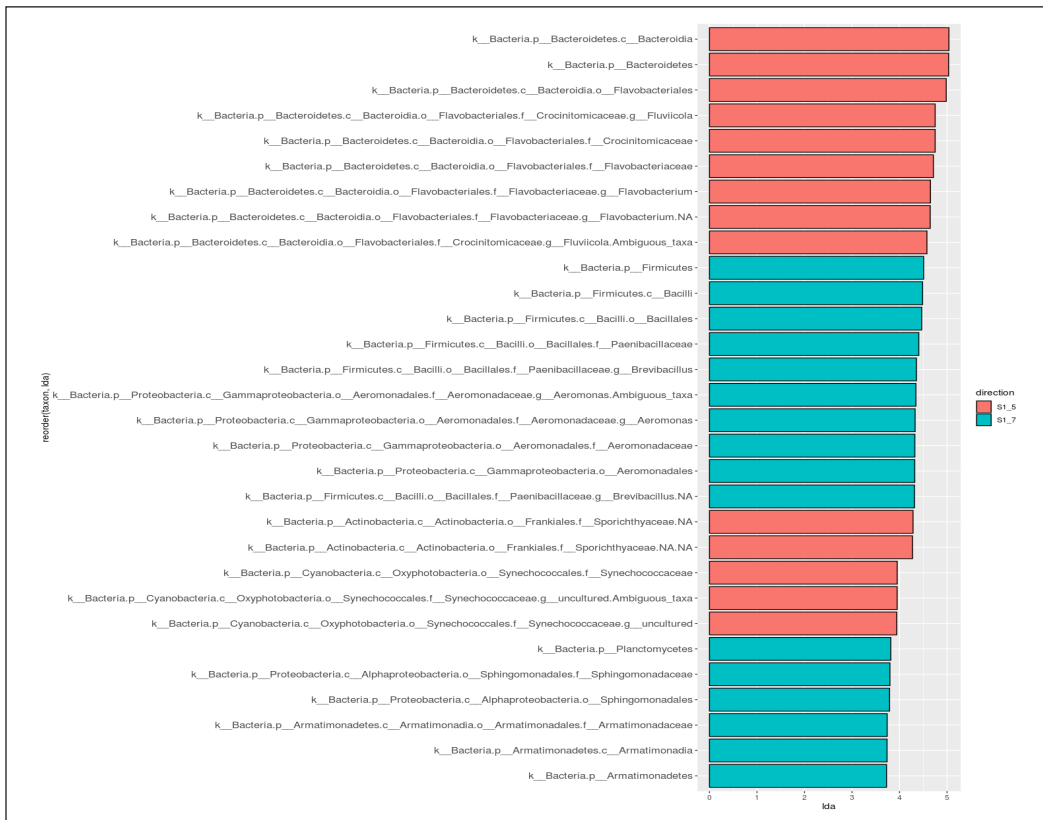
Supplementary Figure 2. Indicator bacterial groups at Site 1 on wet (S1_5) and dry conditions (S1_7), respectively

Nurul Najwa Farihah Mat Lazim, Afiqah Mohamed, Zana Ruhaizat Zana Rudin, Fatimah Md Yusoff, Ikhsan Natrah and Shahrizim Zulkifly

Supplementary Figure 3. Indicator bacterial groups at Site 2 on wet (S2_5) and dry conditions (S2_7), respectively
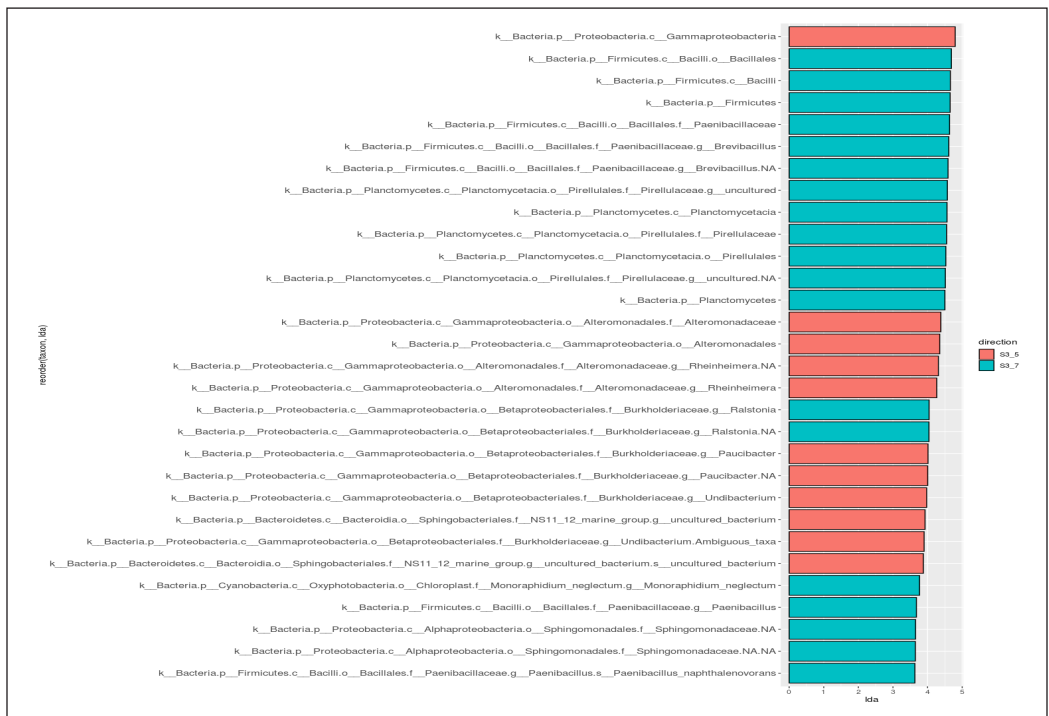
Supplementary Figure 4. Indicator bacterial groups at Site 3 on wet (S3_5) and dry conditions (S3_7), respectively

Supplementary Table 1
List of bacteria identified at the species level

Phylum	Class	Species
Acidobacteria	Acidobacteriia	*Edaphobacter modestum*
Actinobacteria	Actinobacteria	*Candidatus Aquiluna rubra*
		Virgisporangium ochraceum
		Mycobacterium arupense
		Mycobacterium celatum
		Propionibacterium acnes
		Pseudonocardia halophobica
		Bifidobacterium adolescentis
		Bifidobacterium longum
	Coriobacteriia	*Collinsella aerofaciens*
Bacteroidetes	Saprospirae	*Aquirestis calciphila*
	Bacteroidia	*Bacteroides caccae*
		Bacteroides uniformis
		Prevotella copri
	Cytophagia	*Flexibacter ruber*
	Flavobacteriia	*Elizabethkingia meningoseptica*
		Flavobacterium columnare
		Flavobacterium succinicans
	Sphingobacteriia	*Solitalea canadensis*
		Sphingobacterium multivorum
Cyanobacteria	Chloroplast	*Acutodesmus obliquus*
	Nostocophycideae	*Cylindrospermopsis raciborskii*
Firmicutes	Bacilli	*Anoxybacillus kestanbolensis*
		Lactococcus garvieae
	Clostridia	*Faecalibacterium prausnitzii*
Fusobacteria	Fusobacteriia	*Cetobacterium somerae*
Proteobacteria	Alphaproteobacteria	*Asticcacaulis biprosthecium*
		Brevundimonas diminuta
		Ochrobactrum intermedium
		Hyphomicrobium sulfonivorans
		Methylobacterium organophilum
		Paracoccus marcusii
		Rhodovarius lipocyclicus
		Phaeospirillum fulvum
		Orientia tsutsugamushi
		Blastomonas natatoria
		Sphingomonas changbaiensis
		Sphingomonas echinoides
		Sphingomonas wittichii
		Sphingomonas yabuuchiae
	Betaproteobacteria	*Salinispora tropica*
		Acidovorax delafieldii

Supplementary Table 1 *(continue)*

Phylum	Class	Species
		Roseateles depolymerans
		Variovorax paradoxus
		Janthinobacterium lividum
		Polynucleobacter cosmopolitanus
		Methylotenera mobilis
	Deltaproteobacteria	*Bdellovibrio bacteriovorus*
		Desulfovibrio putealis
	Epsilonproteobacteria	*Sulfuricurvum kujiense*
	Gammaproteobacteria	*Plesiomonas shigelloides*
		Acinetobacter rhizosphaerae
		Acinetobacter schindleri
		Pseudomonas nitroreducens
		Nevskia ramosa
		Luteibacter rhizovicinus
		Lysobacter brunescens
		Pseudoxanthomonas mexicana
Spirochaetes	Leptospirae	*Leptospira biflexa*
	Spirochaetes	*Spirochaeta aurantia*
Verrucomicrobia	Verrucomicrobiae	*Akkermansia muciniphila*
		Prosthecobacter debontii

Supplementary Table 2
Top 10 genera normalised abundancy in percentage

	Genus	S1R15	S1R17	S1R25	S1R27	S1R35	S1R37	S2R15	S2R17	S2R25	S2R27	S2R35	S2R37	S3R15	S3R17	S3R25	S3R27	S3R35	S3R37
1	g__hgcI clade	0.018	0.423	0.091	0.373	0.416	0.112	0.047	0.213	0.319	0.226	0.189	0.220	0.194	0.008	0.162	0.011	0.065	0.289
2	g__Flavobacterium	0.536	0.001	0.220	0.001	0.118	0.255	0.396	0.000	0.165	0.005	0.257	0.000	0.259	0.029	0.281	0.001	0.016	0.000
3	g__Cyanobium PCC-6307	0.026	0.170	0.155	0.159	0.118	0.222	0.223	0.173	0.192	0.177	0.141	0.170	0.110	0.109	0.083	0.042	0.166	0.089
4	g__uncultured	0.004	0.002	0.001	0.001	0.001	0.014	0.113	0.165	0.059	0.250	0.050	0.155	0.120	0.352	0.139	0.221	0.153	0.264
5	g__CL500-29 marine group	0.056	0.279	0.039	0.237	0.090	0.143	0.050	0.039	0.063	0.038	0.081	0.037	0.117	0.035	0.116	0.025	0.050	0.133
6	g__Pseudomonas	0.301	0.002	0.024	0.001	0.007	0.057	0.043	0.354	0.000	0.213	0.028	0.361	0.001	0.007	0.025	0.000	0.000	0.003
7	g__uncultured	0.001	0.001	0.000	0.000	0.001	0.002	0.109	0.053	0.120	0.068	0.066	0.055	0.179	0.047	0.174	0.071	0.163	0.184
8	g__Fluviicola	0.057	0.008	0.470	0.007	0.250	0.004	0.019	0.003	0.083	0.005	0.189	0.002	0.021	0.013	0.018	0.005	0.002	0.006
9	g__Brevibacillus	0.000	0.115	0.000	0.221	0.000	0.179	0.000	0.000	0.000	0.011	0.000	0.000	0.000	0.169	0.000	0.347	0.000	0.032
10	g__uncultured	0.001	0.000	0.000	0.000	0.000	0.013	0.000	0.000	0.000	0.007	0.000	0.000	0.001	0.231	0.001	0.277	0.386	0.000