

Article

The Community Structure of eDNA in the Los Angeles River Reveals an Altered Nitrogen Cycle at Impervious Sites

Savanah Senn ^{1,2,*} , Sharmodeep Bhattacharyya ^{1,3}, Gerald Presley ^{1,4}, Anne E. Taylor ⁵, Rayne Stanis ², Kelly Pangell ² , Daila Melendez ^{2,5} and Jillian Ford ²

¹ Environmental Sciences Graduate Program, Oregon State University, Corvallis, OR 97331, USA; sharmodeep.bhattacharyya@oregonstate.edu (S.B.); gerald.presley@oregonstate.edu (G.P.)

² Department of Agriculture Sciences, Los Angeles Pierce College, 6201 Winnetka Avenue, PMB 553, Woodland Hills, CA 91304, USA; raynelace@gmail.com (R.S.); pangellk14875@student.laccd.edu (K.P.); waggond1440@student.laccd.edu (D.M.); fordjm7164@student.laccd.edu (J.F.)

³ Department of Statistics, Oregon State University, Corvallis, OR 97331, USA

⁴ Department of Wood Science & Engineering, Oregon State University, Corvallis, OR 97331, USA

⁵ Department of Crop and Soil Sciences, Oregon State University, Corvallis, OR 97331, USA

* Correspondence: stclais@piercecollege.edu

Abstract: In this study, we sought to investigate the impact of urbanization, the presence of concrete river bottoms, and nutrient pollution on microbial communities along the L.A. River. Six molecular markers were evaluated for the identification of bacteria, plants, fungi, fish, and invertebrates in 90 samples. PCA (principal components analysis) was used along with PAM (partitioning around medoids) clustering to reveal community structure, and an NB (negative binomial) model in DESeq2 was used for differential abundance analysis. PCA and factor analysis exposed the main axes of variation but were sensitive to outliers. The differential abundance of Proteobacteria was associated with soft-bottom sites, and there was an apparent balance in the abundance of bacteria responsible for nitrogen cycling. Nitrogen cycling was explained via ammonia-oxidizing archaea; the complete ammonia oxidizers, *Nitrospira* sp.; nitrate-reducing bacteria, *Marmoricola* sp.; and nitrogen-fixing bacteria *Devosia* sp., which were differentially abundant at soft-bottom sites (p adj < 0.002). In contrast, the differential abundance of several *cyanobacteria* and other anoxygenic phototrophs was associated with the impervious sites, which suggested the accumulation of excess nitrogen. The soft-bottom sites tended to be represented by a differential abundance of aerobes, whereas the concrete-associated species tended to be alkaliphilic, saliniphilic, calciphilic, sulfate dependent, and anaerobic. In the Glendale Narrows, downstream from multiple water reclamation plants, there was a differential abundance of cyanobacteria and algae; however, indicator species for low nutrient environments and ammonia-abundance were also present. There was a differential abundance of ascomycetes associated with Arroyo Seco and a differential abundance of *Scenedesmacene* green algae and cyanobacteria in Maywood, as seen in the analysis that compared suburban with urban river communities. The proportion of Ascomycota to Basidiomycota within the L.A. River differed from the expected proportion based on published worldwide freshwater and river 18S data; the shift in community structure was most likely associated with the extremes of urbanization. This study indicates that extreme urbanization can result in the overrepresentation of cyanobacterial species that could cause reductions in water quality and safety.

Keywords: metabarcoding; microbial communities; nitrogen cycling; statistical modeling; water quality; urban river ecology



Citation: Senn, S.; Bhattacharyya, S.; Presley, G.; Taylor, A.E.; Stanis, R.; Pangell, K.; Melendez, D.; Ford, J. The Community Structure of eDNA in the Los Angeles River Reveals an Altered Nitrogen Cycle at Impervious Sites. *Diversity* **2023**, *15*, 823. <https://doi.org/10.3390/d15070823>

Academic Editors: Alessandro Ferrarini and Manuel Elias-Gutierrez

Received: 19 March 2023

Revised: 15 May 2023

Accepted: 18 May 2023

Published: 29 June 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The Los Angeles River has the potential to influence systems beyond its boundaries, such as estuarine environments at its outlet to the Pacific Ocean. In 2020, the County of Los Angeles' gross domestic product was USD 6.5 billion [1] and its population was

over 10 million [2]. Contamination, such as heavy metals, excess nutrients, coliform bacteria, and cyanide [3], have resulted from industrialization and high population. The L.A. River is a habitat for bacteria, fungi, fish, plants, and invertebrates that are sensitive to pollution. More recently, efforts have focused on protection and recognition of the river as a natural ecosystem, and part of that effort has been assessing the impacts of urbanization on the L.A. River ecosystems through eDNA sampling [4].

There have been few studies which have aimed to characterize the biome of the L.A. River; however, interest in characterizing microbial communities in this biome has increased in recent years [5,6]. The diversity of life, including fungi, bacteria, plants, fish, and invertebrates is indicative of ecosystem health. The presence or absence of certain “indicator” species reflect health and the presence of oxygen or degradation and pollution [7–10]. By investigating microbial community composition and identifying relative species abundance, ecosystem health can be compared among different locations subject to different pollutant profiles.

The L.A. River is unique and the impact of various types of urban pollution and infrastructure on microbial communities may be studied readily. The river runs through rural, suburban, and urban areas and the impact of population density can be assessed. The Los Angeles River was highly modified to facilitate flood control [10–12], due to flooding, which could be catastrophic.

A crucial question documented by Wenger et al. refers to the relationship between urbanization and the structure and function of microbial communities, which has not been well studied [13]. The question of how microbial communities may differ from one another in different land use areas and how urbanization may affect the proportions of different classes of microbes remains vital. The importance of this type of investigation was also underscored in the perspective of Antwis et al. on the most important areas of inquiry in microbial ecology [14]. In terms of urbanization, the modification or toxification of the environment may have influenced which microorganisms were present.

Habitats were diminished due to most of the L.A. River bottom being impervious concrete [15]. One of the key aspects of the paved condition is the decrease in plant life, which would absorb excess nitrogen in the environment through its roots. According to Wenger et al., an inquiry into the characteristics of piped or concrete paved tributaries as they influence biogeochemical processes represents one of the most important topics in urban stream ecology [13]. The presence of a concrete river bottom has been known to influence the oxygen content of freshwater, and this factor is expected to be one of the key factors which influences communities existing under a concrete-bottom condition. Nevertheless, if river organisms, such as oxygenic autotrophs, generate oxygen for the aboveground environment, it would help to offset such a concern as it performs a beneficial function. However, if cyanobacteria dominate, they may generate excess nitrogen which would not be absorbed by plants under paved conditions.

Since bacteria play a huge role in the breakdown of waste, nitrogen cycling, plant growth promotion, and pathogenicity, differences in bacteria warranted a closer look. A lack of oxygen in the underwater environment was expected to be one of the key factors which would influence communities under concrete-bottom conditions. Furthermore, concrete-paved rivers contribute to the urban heat island effect, which involves increased light intensity and heat [13]. Urban rivers generally have a cooling effect on a metropolis by virtue of the water that flows along them and the green spaces they support [16].

In the absence of rain, the L.A. River is fed by water from three water reclamation plants. Ackerman et al. found in 2003 that there were higher ratios of ammonia to nitrate near the water reclamation plants [10]. The benefits of using reclaimed water are obvious in terms of ecosystem services, as a river fed by recycled water would be expected to provide more habitat than a dry riverbed. The year-round supply of water has the potential to support wildlife and vegetation. The water sources have been shown to increase the NO_3^- concentration near the treatment plant sources, but it also would be expected to dilute the concentration of other pollutants, such as hydrocarbons from households and

industry pollutants, such as heavy metals. The proximity to a water reclamation plant could influence the diversity of bacterial sequences recovered from different sampling locations. A potential abundance of nitrate from water treatment plants was a concern historically at sites near Glendale [10]; however, the flow of water to wildlife would be expected to promote diversity. On balance, the river would be a dry ravine for most of the year due to the Mediterranean climate, if it were not for the releases from the water treatment plants.

In this study we sought to investigate the impact of urbanization, the presence of concrete river bottoms, and nutrient pollution on microbial communities along the L.A. River. This was achieved via meta-barcoding and the community analysis of environmental DNA (eDNA). Organisms that perform beneficial functions in the L.A. River ecosystem were identified and quantified from samples taken along the river [17]. This study focused on both eukaryotic and prokaryotic microbes, including archaea, bacteria, cyanobacteria, fungi, plants, and eukaryotic algae. Differences in the abundance of these organism types were measured and analyzed in order to test for statistically significant differences in composition between the sites of interest, i.e., differential abundance. This work contributes to a better understanding of the microbial ecology of the L.A. River ecosystem and helps identify urbanization impacts on microbial communities.

2. Materials and Methods

2.1. Sample Collection

The original data were generated as part of a BioBlitz program by University of California CALeDNA. CALeDNA is a collaboration of scientists creating a baseline of data for the biodiversity in California. Samples were collected by the UC CALeDNA team led by Miroslava Ramos, the project manager. Ninety replicated samples were collected from sediment over a 51-mile span of the channelized portion of the L.A. River and its tributaries. Three subsamples were taken from each sampling location and bulked after DNA extraction to capture a picture of the diversity within a 3-foot radius. In total, there were 180 subsamples.

Table 1 lists the sampling sites by their GPS coordinates for reference. The sampling sites were spread throughout the L.A. River Watershed. Tillman WRP is near the Sepulveda Dam. Note that Verdugo Wash flowed to Glendale Narrows, and Glendale also received water from the intermediary Glendale Water Reclamation Plant. Additionally, depicted is Arroyo Seco, a naturalized area that flows into the industrialized area of Maywood, providing contrast.

2.2. DNA Isolation and Amplification

DNA was extracted using the Qiagen DNEasy PowerSoil Kit. Six molecular markers specific to different kingdoms of life were amplified from the eDNA for amplicon sequencing. Amplicon libraries from each sample type with Illumina barcode adapters were sequenced on the MiSeq platform at 35,000 paired reads each. Quality control was performed in QIIME [18]. Cutadapt was used to remove Illumina adaptor sequences, and DADA2 was used for quality score trimming and the identification of unique ASVs. Taxonomies were assigned to amplicon sequence variants with an 80% likelihood cutoff from the CRUX database. A GreenGenes classifier was used. Each marker dataset was output into an ASV (amplicon sequence variant) table for downstream analysis using the Anacapa toolkit [19]. Table 1 shows the primer used for each marker in the dataset; in Table 2 metadata is provided for each of the samples.

Table 1. Tabulation of the types of genomic data that were available for the L.A. River [20].

Marker	Description	Target Organisms	Forward Primer	Reverse Primer	Reference
FITS	Fungal rRNA Internal Transcribed Spacer	Fungi	GTCGGTAAACTCGTGCCAGC	CATAGTGGGTATCTAATCCCAGTTTG	Yang et al., 2018 [21]
16S	Prokaryotic rRNA small subunit	Bacteria, archaea	GTGYCAGCMGCCGCGTAA	GGACTACNVGGGTWTCTAAT	F: 515F and R: 806R, see Caporaso et al., 2012 [22]
18S	Eukaryotic rRNA small subunit	Fungi, algae, protists	GTACACACCGCCCGTC	TGATCCTTCTGCAGGTTCACCTAC	Amaral-Zettler et al., 2009 [23]; Euk_1391f and EukBr
CO1	Mitochondrial cytochrome oxidase subunit I	Animals	ATGCGATACTGGTGTGAAT	GACGCTTCTCCAGACTACAAT	Gu et al., 2013 [24]
12S	Mitochondrial rRNA small subunit	Fish, birds, snakes, insects	GGWACWGGWTGAACWGTWTAYCCYCC	TANACYTCnGGRTGNCCRAARAAYCA	Leray et al., 2013 [25]
PITS	Plant rRNA Internal Transcribed Spacer	Plants	GGAAGTAAAGTCGTAACAAGG	CAAGAGATCCGTTGTGAAAGTT	F: ITS5, White et al., 1990 [26]; R: 5.8S, Epp et al., 2012 [27]

Table 2. The table of metadata for the L.A. River sites, showing the distribution of the samples across the site features.

Sample No.	LA River Site	Latitude	Longitude	Habitat	River Condition
K0585_T9	Arroyo Seco	34.203154	−118.166402	Frequently submerged, intertidal, marsh	soft
K0593_C3	Arroyo Seco	34.203274	−118.166417	Terrestrial, not submerged	soft
K0594_E4	Arroyo Seco	34.202987	−118.166335	Terrestrial, not submerged	soft
K0595_B2	Arroyo Seco	34.203593	−118.166448	Terrestrial, not submerged	soft
K0595_L7	Arroyo Seco	34.203567	−118.166415	Terrestrial, not submerged	soft
K0595_T9	Arroyo Seco	34.204139	−118.166314	Terrestrial, not submerged	soft
K0597_M8	Arroyo Seco	34.20375	−118.166481	Terrestrial, not submerged	soft
K0599_L7	Arroyo Seco	34.20331	−118.166408	Frequently submerged, intertidal, marsh	soft
K0526_B2	Bowtie Parcel	34.108161	−118.246186	Fully submerged	soft
K0529_L7	Bowtie Parcel	34.108149	−118.246176	Fully submerged	soft
K0672_C3	Bowtie Parcel	34.108433	−118.246959	Fully submerged	soft
K0672_G5	Bowtie Parcel	34.108278	−118.246926	Fully submerged	soft
K0674_E4	Bowtie Parcel	34.108186	−118.246584	Fully submerged	soft
K0678_E4	Bowtie Parcel	34.108131	−118.246003	Fully submerged	soft
K0679_B2	Bowtie Parcel	34.108278	−118.246341	Fully submerged	soft
K0679_M8	Bowtie Parcel	34.108374	−118.246774	Fully submerged	soft
K0528_A1	Bull Creek	34.181558	−118.497717	Frequently submerged, intertidal, marsh	soft
K0528_E4	Bull Creek	34.182029	−118.49771	Frequently submerged, intertidal, marsh	soft
K0528_K6	Bull Creek	34.181975	−118.497849	Frequently submerged, intertidal, marsh	soft
K0529_K6	Bull Creek	34.181652	−118.497718	Frequently submerged, intertidal, marsh	soft
K0529_T9	Bull Creek	34.181651	−118.497716	Fully submerged	soft

Table 2. Cont.

Sample No.	LA River Site	Latitude	Longitude	Habitat	River Condition
K0530_A1	Bull Creek	34.181419	−118.497763	Frequently submerged, intertidal, marsh	soft
K0530_B2	Bull Creek	34.181342	−118.497657	Frequently submerged, intertidal, marsh	soft
K0530_E4	Bull Creek	34.1814	−118.497865	Frequently submerged, intertidal, marsh	soft
K0528_G5	Compton Creek	33.843656	−118.206466	Frequently submerged, intertidal, marsh	soft
K0528_L7	Compton Creek	33.843055	−118.205667	Fully submerged	soft
K0528_T9	Compton Creek	33.843328	−118.2061	Frequently submerged, intertidal, marsh	soft
K0529_A1	Compton Creek	33.843196	−118.205854	Frequently submerged, intertidal, marsh	soft
K0530_C3	Compton Creek	33.843311	−118.206092	Frequently submerged, intertidal, marsh	soft
K0530_K6	Compton Creek	33.842877	−118.205544	Frequently submerged, intertidal, marsh	soft
K0530_L7	Compton Creek	33.842749	−118.205402	Fully submerged	soft
K0530_M8	Compton Creek	33.843196	−118.205854	Frequently submerged, intertidal, marsh	soft
K0529_C3	Elysian Valley	34.083829	−118.228152	Fully submerged	concrete
K0672_T9	Elysian Valley	34.084621	−118.228071	Frequently submerged, intertidal, marsh	concrete
K0673_A1	Elysian Valley	34.084217	−118.228066	Frequently submerged, intertidal, marsh	concrete
K0673_G5	Elysian Valley	34.084227	−118.228048	Fully submerged	concrete
K0674_G5	Elysian Valley	34.08455	−118.228053	Fully submerged	concrete
K0676_B2	Elysian Valley	34.08449	−118.228157	Fully submerged	concrete
K0676_T9	Elysian Valley	34.084721	−118.228145	Fully submerged	concrete
K0677_A1	Elysian Valley	34.084482	−118.228157	Frequently submerged, intertidal, marsh	concrete
K0593_T9	Glendale	34.155282	−118.275211	Fully submerged	concrete
K0594_L7	Glendale	34.15459	−118.276618	Fully submerged	concrete
K0596_C3	Glendale	34.155107	−118.275459	Fully submerged	concrete
K0596_E4	Glendale	34.154774	−118.27637	Frequently submerged, intertidal, mars	concrete
K0596_L7	Glendale	34.154918	−118.276231	Fully submerged	concrete
K0596_T9	Glendale	34.154973	−118.275799	Fully submerged	concrete
K0597_K6	Glendale	34.154997	−118.275944	Fully submerged	concrete
K0597_L7	Glendale	34.155157	−118.27542	Fully submerged	concrete
K0526_C3	Glendale Narrows	34.102813	−118.242742	Fully submerged	concrete
K0526_G5	Glendale Narrows	34.103427	−118.242642	Fully submerged	concrete
K0529_B2	Glendale Narrows	34.103109	−118.242634	Fully submerged	soft
K0529_G5	Glendale Narrows	34.103652	−118.242686	Fully submerged	concrete

Table 2. Cont.

Sample No.	LA River Site	Latitude	Longitude	Habitat	River Condition
K0529_M8	Glendale Narrows	34.103251	−118.242645	Fully submerged	concrete
K0672_B2	Glendale Narrows	34.10274	−118.242669	Fully submerged	concrete
K0678_B2	Glendale Narrows	34.103274	−118.242544	Fully submerged	concrete
K0678_K6	Glendale Narrows	34.103437	−118.24275	Fully submerged	concrete
K0672_A1	Long Beach	33.762909	−118.202355	Fully submerged	soft
K0674_M8	Long Beach	33.762738	−118.202271	Fully submerged	concrete
K0676_M8	Long Beach	33.762683	−118.202126	Fully submerged	concrete
K0677_B2	Long Beach	33.762833	−118.202418	Fully submerged	concrete
K0677_E4	Long Beach	33.762907	−118.202298	Fully submerged	concrete
K0677_L7	Long Beach	33.762841	−118.20235	Fully submerged	concrete
K0678_L7	Long Beach	33.762906	−118.202305	Fully submerged	soft
K0701_C3	Long Beach	33.76269	−118.202303	Fully submerged	concrete
K0527_A1	Maywood	33.986755	−118.171412	Frequently submerged, intertidal, marsh	concrete
K0527_C3	Maywood	33.988033	−118.172607	Fully submerged	concrete
K0527_E4	Maywood	33.987023	−118.171842	Fully submerged	concrete
K0527_K6	Maywood	33.986686	−118.171342	Fully submerged	concrete
K0527_L7	Maywood	33.987668	−118.172288	Fully submerged	concrete
K0527_T9	Maywood	33.986617	−118.171324	Fully submerged	concrete
K0539_L7	Maywood	33.986776	−118.17165	Fully submerged	concrete
K0593_G5	Sepulveda Dam	34.168961	−118.475292	Fully submerged	soft
K0594_A1	Sepulveda Dam	34.168698	−118.475195	Fully submerged	soft
K0594_T9	Sepulveda Dam	34.168961	−118.475292	Fully submerged	soft
K0595_G5	Sepulveda Dam	34.168941	−118.47461	Terrestrial, not submerged	soft
K0597_T9	Sepulveda Dam	34.1688	−118.475049	Fully submerged	soft
K0599_G5	Sepulveda Dam	34.16868	−118.474846	Frequently submerged, intertidal, marsh	soft
K0599_K6	Sepulveda Dam	34.168906	−118.475125	Fully submerged	soft
K0599_T9	Sepulveda Dam	34.168758	−118.474733	Rarely submerged, wetland, arroyo	soft
K0593_A1	Tujunga Wash	34.258032	−118.386781	Fully submerged	concrete
K0593_E4	Tujunga Wash	34.258403	−118.386614	Fully submerged	concrete
K0595_M8	Tujunga Wash	34.257481	−118.386845	Fully submerged	concrete
K0596_B2	Tujunga Wash	34.258667	−118.386473	Fully submerged	concrete
K0597_E4	Tujunga Wash	34.258716	−118.386376	Fully submerged	concrete
K0599_A1	Tujunga Wash	34.258424	−118.386387	Fully submerged	concrete
K0599_E4	Tujunga Wash	34.258395	−118.386592	Fully submerged	concrete
K0599_M8	Tujunga Wash	34.258016	−118.386744	Fully submerged	concrete
K0593_L7	Verdugo Wash	34.203216	−118.237654	Fully submerged	soft
K0595_A1	Verdugo Wash	34.202985	−118.237755	Fully submerged	soft
K0596_G5	Verdugo Wash	34.202611	−118.237615	Fully submerged	soft

For this differential abundance analysis, computation focused on the bacteria and fungi. However, the results of the differential abundance analysis may also include algae and nematodes, for example. Table 3 shows the covariates that were contrasted in DESeq2.

Table 3. List of covariates that were tested for association with a differential abundance of bacterial and fungal taxa.

Marker	Covariate	Factor Levels Tested
16S	LA River Site	Glendale Narrows, Verdugo Wash
16S	River Condition	Soft-Bottom, Concrete
16S	Habitat	Frequently Submerged, Fully Submerged
FITS	Habitat	Frequently Submerged, Fully Submerged
FITS	LA River Site	Maywood, Arroyo Seco

2.3. Statistical Approach

The goal of this project was to examine sample diversity using a variety of methods using a Euclidean distance matrix [28]. The Euclidean distance is given by [29]:

$$d(j_1, j_2) = [(X1j_1 - X1j_2)^2 + \dots + (Xnj_1 - Xnj_2)^2]^{1/2}$$

The methods utilizing the Euclidean dissimilarity measure will include the neighbor joining of samples [30], the UPGMA of samples [30], heatmap visualization using the chi-square standardization of samples, and PAM (partitioning around medoids) clustering applied to PCA. Ranacapa [31] was used to perform a PERMANOVA beta diversity test and visualize with principal coordinates analysis (PCoA) to help with hypothesis development.

PAM clustering was applied to PCA to investigate whether samples cluster by location in an unsupervised model and whether the PCA reflected a spatial relationship inherent in the genetic distances. The PAM function from the cluster package was used [32]: First, K representative medoids are arbitrarily selected, then swapping cost C_{ih} to swap medoid h and non-medoid i is calculated. If the resulting value is negative, then the medoid and non-medoid are swapped. The process is repeated until there is no change. Principal components analysis reveals population stratification and PAM is used for classification of samples.

The classification of samples was expected based on the taxonomic composition of samples; that is, if there were differentially abundant taxa between groupings then separation into different PAM clusters would be expected. To select the optimal number of clusters K , the PAM model with the highest average silhouette value was selected. The factor analysis of the most important taxon features in the PCA for each marker dataset gave some preliminary evidence about which particular taxa may be differentially abundant. Relative abundance was compared for important plant taxa using a pivot table in Microsoft Excel.

2.4. Chi Square Test of Proportions for the 18S Marker

The data were published originally as “Table 2, Richness of Main Taxonomic Groups of Fungi in Freshwater Ecosystems” from a study that has count data for the main taxonomic groups of fungi in freshwater ecosystems that can be used as a comparison [33]. The information captures data from 22 publicly available datasets from around the world. The initial exploration of the data revealed that there were few Cryptomycota and Chytridiomycota identified in the pooled L.A. River samples. The chi-square test tested whether the proportion of Ascomycota:Basidiomycota in the L.A. River differed significantly from that of freshwater and river environments in the published data. The hypotheses that were tested for this analysis are contained in Supplemental Materials.

Overdispersion is common in taxonomic count data for environmental samples. The model that was implemented in DESeq2 to answer these research questions was a negative

binomial model. In these data, zero inflation is also suspected. The way that DESeq2 dealt with overinflation in this analysis was to analyze only positive counts. Exploratory plots for dispersion in the fungi dataset were generated to further investigate the appropriateness of the model (see Supplemental Materials).

2.5. Differential Abundance Analysis

For differential abundance analysis, DESeq2 was employed [34]. The DESeq2 package has handled RNA-seq or ChIP-seq, metabarcoding ASV tables, and any similar genomic data that consisted of counts. The goal was to correct some problems associated with using chi-square test and the Poisson distribution for these types of data, which may not effectively control a Type I error [34].

It was assumed that the number of reads in sample j assigned to gene or taxon $i = K_{ij} \sim NB(\mu_{ij}, \sigma^2)$ follows a negative binomial distribution (NB), which is commonly used for the modelling of data in the presence of overdispersion [34].

The following further assumptions were made:

1. The mean parameter is the expectation value for K_{ij} and is proportional to the actual number of sequence counts for gene i under the experimental condition p . The size factor is also accounted for, which is essentially the coverage or sequencing depth of the genetic library for each sample.
2. The variance σ^2 is the sum of the shot noise and the raw variance.
3. The model uses a pooled variance from genes (or taxa) with similar count values to estimate the per gene raw variance.

K_{ij} follows a Poisson distribution. If the rate that fragments are assigned to known sequences depends on a random variable $R_{ij} = r_{ij}$, and the size factor, s_{ij} , then when R_{ij} is modeled by the gamma distribution, $K_{ij} \sim NB(\mu_{ij}, \sigma^2)$, the cycle has been completed.

In terms of fitting the model, data exist in a $n \times m$ table of K_{ij} counts: $i = 1 \dots n$ genes in $j = 1 \dots m$ samples. The parameters used were:

1. m size factors, including 1 for each sample.
2. n expression strength parameters q_{ip} for each condition p . In other words, the expectation values for the abundance of counts for gene or taxon i are proportional to q_{ip} .
3. The pooled variance parameter simulates the dependence of V_{ip} on the expectation value for the mean, q_{ip} , for each condition p .

The size factor s_{ij} allows comparisons between samples with different sequencing depths. Size factors are estimated via the median of observed count ratios [34]. q_{ip} is estimated through a transformation of the average counts from j samples under condition p . The fit can be applied to small numbers of replicates using local regression to estimate the raw variance. The method is a gamma family GLM for a local regression that implements R locfit.

A hypothesis rejection in DESeq would mean that the difference in counts between two samples was larger than would be expected if the samples were replicates from the same individual or tissue [34]; the rejection does not indicate what is responsible for the difference. A rejection shows that a taxon, protein, or gene count was differentially abundant between two samples. However, a hypothesis rejection would not reveal if it was more different than what would typically be seen if two separate locations along the same river were sampled. It would also not reveal whether the difference would have a greater magnitude than if one compared the differential abundance of that taxa between two different rivers. It empowers the user to detect differences, while controlling the Type I error. Volcano plots were subsequently visualized in SystemPipeR [35] and Enhanced Volcano [36].

3. Results

The Unweighted Unifrac distance method coupled with PERMANOVA, visualized by PrinCoA, was the most sensitive for the detection of differences between groups based on sampling site, habitat, or depth. The chi-squared standardized heatmap was not sensitive.

PCA alone was not sensitive, although the factor loadings were useful for revealing the few important taxa that differed between samples. PAM coupled with PCA was more useful for identifying highly similar groups of samples and elucidating community structure. PCA with PAM gave a better visualization than the hierarchical clustering methods for this sample size, although overall, the PAM and UPGMA results were very similar.

Table 3 shows the medians and ranges for taxon abundance and sequences per sample. The FITS marker had a median number of sequences per sample of 18,157. Table S1 displays the summary statistics resulting from the NJ (neighbor joining) and UPGMA (unweighted pair group method with arithmetic mean) tree analyses in R phyloseq. As shown in Table S1, the branch length means were similar, but the variance is higher for neighbor joining, with respect to the FITS marker. A higher variance for neighbor joining would be expected.

Depicted in Figure S3 is the PCA for the fungal ITS sequences that were recovered from the L.A. River sediment samples. The first two principal components capture about 37% of the variation in the data. Fungi samples separate high on PC 2 based on the abundance of *Penicillium*, which may be important for the decomposition of leaf litter along the river, and *Cladosporium* sequences, which produce the antibiotic and antimalarial metabolite Cladosporin [37]. Low on PC 2, the separation is based on the abundances of the *Desmodesmus armatus* and *Desmodesmus* sp. variants of algae, especially in Maywood, Glendale Narrows, Glendale, and Elysian Valley. These genera have been known to break down radioactive materials.

Other results from the DESeq2 analysis, showed that in frequently submerged river condition samples, there was a significantly higher abundance of fungi and less bacteria, when compared with submerged samples. The volcano plot showing the large number of significant results for fungi based on the FITS marker is visualized in Figure 1. In frequently submerged sediment samples, *Capnioidales* sp. were differentially abundant based on the adjusted p -values ($p < 1 \times 10^{-13}$), as well as *Penicillium* sp. ($p < 0.0005$). Notably, *Tricladium angulatum* ($p < 1.5 \times 10^{-46}$), *Monocillium tenue* ($p < 2.5 \times 10^{-39}$), *Acremonium nepalense* ($p < 5 \times 10^{-30}$), and *Peziza badia* ($p = 9.5 \times 10^{-15}$) were also significantly more abundant in frequently submerged samples.

As shown in Table 3, the 16S assay had a strong median number of sequences per sample at 15,178. This shows that the sample had a good sequencing depth. As shown in Table 4, the branch length means are similar but the variance is about 50,000 units higher for neighbor joining, with respect to the 16S marker. The rooted and unrooted trees both indicated $k = 5$ for the number of clusters in the community of bacteria.

Table 4. Summary statistics from the neighbor joining and UPGMA trees for each marker. The trees were generated from the Euclidean distance matrix. The tree topological distances have been provided in the far-right column.

L.A. RIVER	Branch Length NJ		Branch Length UPGMA		NJ vs. UPGMA
Marker	Mean	Variance	Mean	Variance	Tree Distance
FITS	1657	5,419,114	1585	4,124,851	8195
16S	620	460,349	609	417,224	2473
18S	2018	5,534,355	1978	4,278,736	10,919
COI	2312	8,746,132	2114	6,010,691	9697
12S	634	4,710,694	1585	4,124,851	12,130
PITS	1457	6,728,373	1351	4,241,554	8516

Figure S5 shows the PCA for the bacterial 16S DNA sequences that were recovered from the L.A. River sediment samples. The first two principal components capture around 42% of the variation in the data. Bacteria DNA samples are separated by numerous

important taxa factor loadings, such as the abundance of Erythrobacteraceae, Proteobacteria, and *Oscillatoriales cyanobacterium*.

DESeq2 results for FITS

Differential Abundance: Maywood, Arroyo Seco

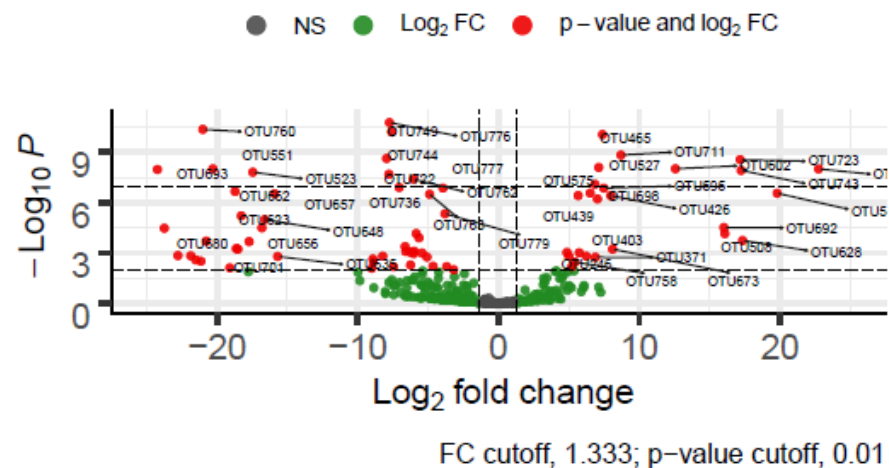


Figure 1. The volcano plot demonstrated the large number of taxa that were differentially abundant between Maywood and Arroyo Seco, with regard to fungi. The differential abundance analysis in DESeq2 for the FITS marker yielded a large number of interesting fungi associated with one location or another (101 significant taxa were detected). NS = not significant, FC = fold change. The OTUs with negative log fold change values were more abundant in Maywood; the OTUs with positive log fold change values were more abundant in Arroyo Seco.

Among others, samples from Maywood and Glendale scored low on PC 2 in terms of high cyanobacteria abundance. Figure 2 shows that the PCA plot for the 16S samples was color coded, corresponding to the best PAM clustering. The best PAM clustering in this case was $k = 4$ with the highest average silhouette width. The samples in the second cluster, colored red, are from Glendale Narrows. The third cluster, colored green, is mostly made up of sediment samples from Maywood and Glendale. The blue and black clusters are made up of a mixture of the remaining sites.

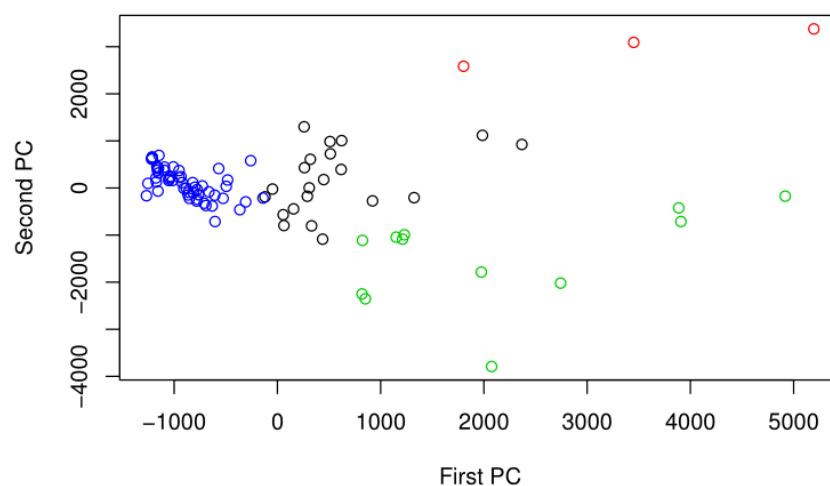


Figure 2. PCA for bacterial identified sequences from the 16S marker by sample, color coded by the best PAM clustering. Each point represents a sample. Note that there is evidence of overdispersion, in particular, high on PC1. The points with the same color were classified by PAM as belonging to the same cluster.

Among the bacteria with a differentially higher abundance of 16S sequences in Glendale Narrows, *Cyanobacteria microcystis* ($p < 1.5 \times 10^{-7}$) and *Oscillatoriales cyanobacterium* ($p < 3 \times 10^{-14}$). *Verrucomicrobia* were also differentially more abundant in Glendale Narrows ($p < 4 \times 10^{-23}$). On the other hand, the alphaproteobacteria *Devosia* from Rhizobiales had differentially higher counts of sequences in samples from Verdugo Wash. These clusters helped inform the DESeq Analysis for Glendale vs. Verdugo Wash and soft-bottom vs. concrete contrasts, the results of which are shown in Tables 5 and 6.

Table 5. The results of the differential abundance analysis for Glendale vs. Verdugo Wash. Positive log fold change results represent sequences that were differentially abundant at the Glendale site. Negative log fold changes represent sequences that were differentially abundant at the Verdugo Wash site.

Taxon	Log ₂ Fold Change	<i>p</i> -adj	Ecological or Metabolic Function and Pathogenicity
<i>Prostheobacter</i> sp.	22.09927	3.71×10^{-23}	possible pathogen, anaerobic, tubulin-like genes, low nutrient environments
<i>Dechloromonas</i> sp.	34.31956	1.53×10^{-41}	may oxidize benzene
<i>Devosia</i> sp.	−22.258	5.73×10^{-5}	nitrogen fixer
<i>Bacillus</i> sp.	−25.3115	1.67×10^{-5}	many beneficial species
Chromatiaceae (unclassified)	23.78784	1.22×10^{-6}	purple sulfur bacteria, use sulfide to fix carbon and generate oxygen
<i>Sandaracinobacter</i> sp.	−30.519	0.009416	metabolism of sulfide to cysteine (or from serine)
<i>Chloroflexaceae</i> (unclassified)	25.68591	0.000938	green non-sulfur bacteria, many heat-loving anoxygenic photoheterotrophs [38,39]
endosymbiont of <i>Ridgeia piscesae</i>	−22.3636	0.00014	gammaproteobacterium, symbiont of a tubeworm
anaerobic bacterium MO-CFX2 <i>Chloroflexi</i>	−6.85917	4.08×10^{-6}	
Rhodocyclales (unclassified)	17.1087	4.15×10^{-8}	nitrogen fixing or nitrogen reducing
<i>Phormidium setchellianum</i>	33.82601	2.58×10^{-14}	potential cause of gastroenteritis, concentrates caused neuro- and hepato-toxicity in mice [40]
<i>Cytophaga xylanolytica</i>	20.18264	0.000268	xylan degrading, does well in sulfogenic and methanogenic environments, anaerobic and gliding
<i>Synechococcus</i> sp.	−23.4117	0.002659	photolysis of sulfide or water, produces neurotoxins [41]
<i>Scenedesmaceae</i> (unclassified)	11.0032	0.000123	green algae, may degrade radioactive materials
<i>Flavobacterium</i> sp.	8.245038	0.000199	often associated with plant resistance to pathogens
<i>Oscillatoriales cyanobacterium</i> HF1	7.271474	0.005122	cyanobacterium which may cause illness or death in humans and animals
<i>Tetrademus obliquus</i>	10.11933	0.001645	produces valuable saturated and unsaturated esters, extract has anticancer and antimicrobial effects [42,43]
<i>Microcystis</i> sp.	28.7773	1.03×10^{-7}	cyanobacterium which is toxic to humans [44]
<i>Rhodocyclaceae</i> bacterium enrichment culture clone Y62	28.91261	5.24×10^{-5}	nitrogen fixing or nitrogen reducing

Table 6. Positive log fold change results represent sequences that were differentially abundant at the soft-bottom sites. Negative log fold changes represent sequences that were differentially abundant at the concrete sites.

Taxon	Log ₂ Fold Change	p-adj	Ecological or Metabolic Function and Pathogenicity
<i>Oscillatoriales cyanobacterium</i> YACCYB599	−25.207183	3.06×10^{-23}	cyanobacteria, which may cause illness or death in humans and animals
<i>Chroococcus subviolaceus</i>	−24.66764915	4.55×10^{-23}	freshwater or high salinity environments, cyanobacteria which can survive with low O ₂ [45]
<i>Haliae</i> sp.	−24.50212313	4.55×10^{-23}	marine gamma proteobacterium, which tolerates up to 12% salinity [46,47]
<i>Halomonas</i> sp.	24.49667323	3.81×10^{-31}	chloride and saline tolerance
<i>Marmoricola</i> sp.	24.12963073	1.43×10^{-27}	denitrifying bacteria [48]
Alpha proteobacterium LS7-MT	10.00393321	8.21×10^{-09}	methanol oxidizer, lives in high temperatures [49]
<i>Nitrosarchaeum koreense</i>	9.188395232	2.37×10^{-18}	aerobic ammonia-oxidizing archaea [50]
Microcystaceae (unclassified)	−8.382519826	0.001244	common eutrophic bloomer, toxin-producing cyanobacterium
<i>Acidobacterium</i> sp. SCGC AAA007-P13	7.849119335	3.12×10^{-7}	potential saprobe
<i>Oscillatoriales cyanobacterium</i> IRH12	−7.732408042	4.32×10^{-8}	cyanobacterium, which may cause illness or death in humans and animals
<i>Roseisolibacter agri</i>	−7.389766623	0.000539	grows in low oxygen environments [51]
<i>Pleurocapsa concharum</i>	−7.310779292	1.03×10^{-7}	ostracod-dependent cyanobacterium [52]
<i>Devosia</i> sp.	7.242636088	5.51×10^{-7}	nitrogen-fixing bacteria
<i>Nitrospira</i> sp. enrichment culture clone LD3	6.970043209	0.001616	nitrifying bacteria, nitrite-oxidizing bacteria
Gamma proteobacterium SCGC AAA007-P21	6.533527317	1.83×10^{-13}	uncultivated bacterioplankton
alpha proteobacterium Schreyahn_AOB_Aster_Kultur_5	6.503508981	0.001529	cultured alphaproteobacterium
Chlamydomonadales (unclassified)	−6.479686479	0.000178	green algae [53]
<i>Chloronema giganteum</i>	−6.382235759	0.000425	photoautotrophic, anoxygenic green non-sulfur bacteria [54]
<i>Chamaesiphon</i> sp.	−6.230017507	0.002384	widely distributed cyanobacterium [55]
<i>Altererythrobacter</i> sp.	6.02052523	0.007591	alkaline or salt tolerant aerobic phototroph, anoxygenic [56–58]
Mycobacteriaceae (unclassified)	5.990283542	0.000524	potential human and animal pathogens
Acidobacteriaceae (unclassified)	5.737312813	2.78×10^{-6}	likely saprobe of plant organic matter
<i>Candidatus</i> <i>Viridilinea mediisalina</i>	−5.72085055	0.009826	anaerobic phototroph, salt-tolerant and prefers alkaline environments [59]
Veillonellaceae bacterium 6–15	−5.56037325	2.59×10^{-5}	bacterial vaginosis
<i>Phormidium setchellianum</i>	−5.548460876	0.000699	cyanobacterium with possible antitumor agents, neuro and hepatotoxicity

Table 6. Cont.

Taxon	Log ₂ Fold Change	p-adj	Ecological or Metabolic Function and Pathogenicity
<i>Calothrix</i> sp. UAM 374	−5.531306605	0.003193	cyanobacterium, which grows on plants and hard substrates [60]
Candidatus <i>Nitrosocosmicus</i> sp.	5.344610141	0.0001	aerobic ammonia-oxidizing archaea
<i>Treponema stenostreptum</i>	−5.019693824	0.003193	syphilis relative
Leptolyngbyaceae (unclassified)	−4.952937198	0.001067	thermophilic and potentially iron-loving cyanobacterium [61]
Holophagaceae (unclassified)	−4.934291389	0.000964	anaerobic dweller of freshwater sediments [62]
Xanthomonadaceae bacterium	−4.711954167	0.002384	potential phytopathogens
<i>Leptolyngbya geysericola</i>	−4.711366069	0.005914	alkaline tolerant non-heterotrophic cyanobacterium, produces calcite on microplastics [63]
Caldilineales bacterium	4.50039412	4.71×10^{-6}	thermophilic and anaerobic [64]
<i>Fusibacter</i> sp. enrichment culture	−4.35065315	0.009823	thiosulfate reducing, potentially halotolerant
<i>Desulfomicrobium</i> sp.	−4.16646108	0.002439	oxidizes sulfide and arsenate in the presence of CO ₂ and acetate [65], reduces nitrate to ammonium [66]
Oscillochloridaceae (unclassified)	−3.874861377	0.005914	anoxygenic phototrophic bacteria [38,67]
Pleurocapsales (unclassified)	−3.695598612	0.009826	cyanobacterium from calcareous environments
<i>Vicinamibacter silvestris</i>	3.602101991	0.002384	polyphosphate accumulating organisms
Firmicutes (unclassified)	2.378738101	0.004923	high abundance in suburban rivers, negatively correlated with ammonia concentration
<i>Stenotrophobacter terrae</i>	2.253024076	0.008829	opportunistic pathogen
Vicinamibacteraceae (unclassified)	2.126473277	0.00044	degrades chitin [68]
Actinobacteria (unclassified)	2.033767588	0.003193	many denitrifying bacteria [69,70]

The soft-bottom river condition was associated with a differentially higher abundance of Alphaproteobacteria and a decreased abundance of *Cyanobacteria pleurocaps* ($p < 1 \times 10^{-6}$) and *Phormidium* ($p < 0.0007$), *Oscillatoria* ($p < 3 \times 10^{-23}$), and *Chroococci* ($p < 5 \times 10^{-23}$) when contrasted with concrete sites. Notably, *Devosia* was more abundant in soft bottoms ($p < 6 \times 10^{-7}$), whereas *Desulfomicrobium* ($p < 0.003$) was more abundant under concrete-bottom conditions. On the other hand, Verrucomicrobia and Haliaceae family Proteobacteria were differentially abundant under soft-bottom conditions ($p < 5 \times 10^{-23}$, $p = 0.01$, respectively).

Most of the bacteria that were differentially expressed in the concrete sites were cyanobacteria and autotrophs. There was also a trend toward a differentially high abundance of DNA sequences from potential human and plant pathogens, including the potential plant pathogen *Xanthomonas*, Clostridia, and bacteria related to the agents that cause reproductive infections. Nevertheless, the soft-bottom sites also had differentially high abundances of Norcardiaceae and Verrucomicrobia, which are also potential pathogens. For the concrete sites, there was a less clear picture of the nitrogen cycle when considering the bacteria alone. There was a clear picture of the nitrogen cycle for the soft-bottom sites, as well as a candidate species for phosphate accumulation.

The highest number of assigned sequences per sample was for the 18S marker, as shown in Table 2. This suggests that the highest overall sequencing depth was for the 18S assay. As shown in Table 4, the branch length means were both near 2000 but the variance was around 125,000 units higher for neighbor joining, with respect to the 18S marker. For both tree topologies, $k = 4$ is apparent for the number of clusters in terms of 18S sequences identified by the assay.

In Figure S6, the PCA for the 18S DNA sequences that were recovered from the L.A. River sediment samples is shown. The first two principal components capture around 46% of the variation in the data. The PCA by sample for 18S validates the FITS results, because the samples scored low on PC 2 based on factor loadings for *Desmodesmus* and other *Scenedesmeceae* taxa of algae. Further, samples scored high on PC 2 based on the *Podocopida* and *Cypridida* high relative sequence abundance. *Podocopida* is a crustacean that comprises freshwater and brine-dwelling groups [71]. The *Cyprididae* are a group of freshwater Ostracods [72]. Figure S7 shows the 18S PCA color coded by the best PAM clustering, which was $k = 5$, with the highest average silhouette width. The red samples in cluster 2 were all from Glendale. Cluster 5, in light blue, corresponds to the Long Beach sediment samples. Considering the spatial heterogeneity displayed by the samples, there is a sense that the genetic material is funneling into Long Beach, reflecting the physical landscape. The fourth cluster, in dark blue, is composed of Sepulveda Dam, Tujunga Wash, and Arroyo Seco.

The observed alpha diversity for fungi sequences based on the 18S marker is shown in Figure 3. Los Angeles River proportions of Ascomycota and Basidiomycota were compared to freshwater and river habitats worldwide. The equality of these proportions were tested on a chi-square distribution. The results showed that the proportions of Ascomycota and Basidiomycota in the L.A. River differed significantly from freshwater and river environments worldwide, based on published 18S data [33].

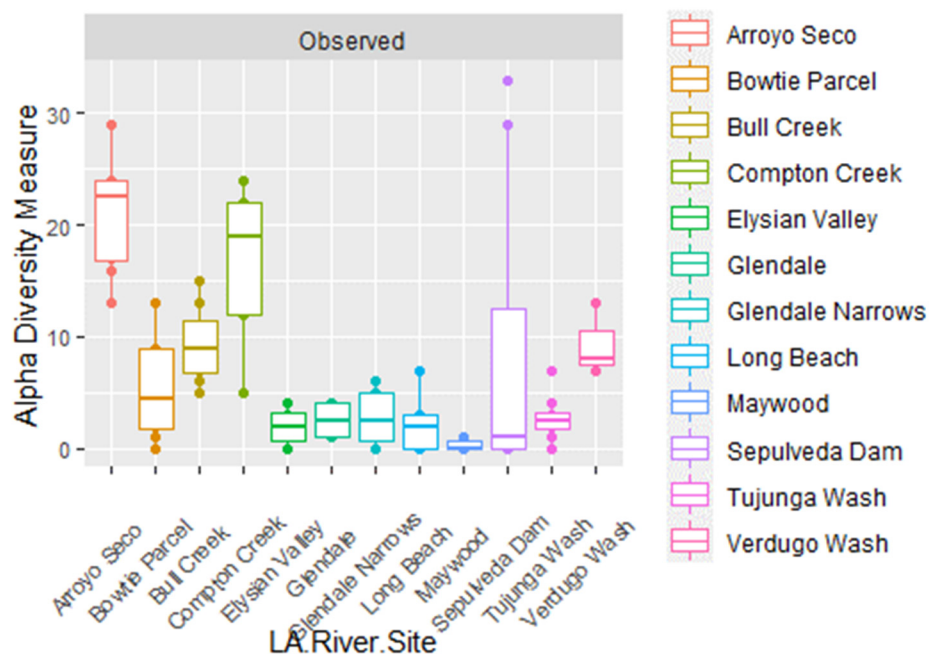


Figure 3. The boxplot of observed alpha diversity shows that the species richness for Ascomycota is the highest in Arroyo Seco, Bull Creek, Compton Creek, and Maywood.

The data that were used for this part of the analysis are publicly available [33] as amplicon sequence variants tables, also known as ASVs or OTUs. OTU stands for operational taxonomic unit. Essentially, these tables have the counts of sequences that were identified from organisms in the environment. The goal is to compare the proportions of different divisions of fungi in the L.A. River to other environments.

In Figure 4, the mosaic plot for the chi-square test of proportions for river habitats worldwide versus those of the L.A. River is shown. The values for the Ascomycota and Basidiomycota in the L.A. River and in worldwide river habitats display a gap between them. This shows that these proportions differ significantly from what one expects if they belonged to the same population. The results of the chi-square test for the equality of proportions shows that the values of Ascomycota and Basidiomycota for the L.A. River are not equal to the proportions of Ascomycota and Basidiomycota in freshwater habitats ($p < 0.0005$) or river habitats ($p < 1 \times 10^{-11}$) described in Lepère's analysis of worldwide freshwater data. In terms of the river habitats, the proportion of Ascomycota to Basidiomycota is 21.5–39.2% higher in the L.A. River. Furthermore, for the freshwater habitat comparisons, the proportion of Ascomycota to Basidiomycota is between 7.3–25.74% higher for the L.A. River, based on the 95% confidence intervals. When comparing the mosaic plots in Figure S9 and Figure 4, the gap between the values of Ascomycota and those of Basidiomycota appears smaller for the L.A. River compared to freshwater habitats in the study by Lepère et al. [33], compared with river environments.

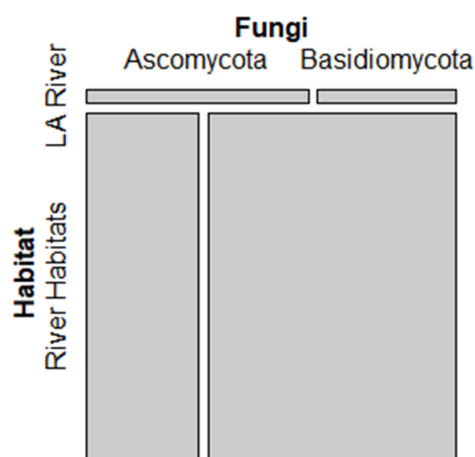


Figure 4. The mosaic plot shows that there is a difference in the proportion of Ascomycetes to Basidiomycetes in the L.A. River compared to river habitats worldwide [33]. This gap was larger than the gap shown in Figure S9 for freshwater habitats.

The alpha diversity analysis for Ascomycetes is plotted in Figure 3. The mosaic plot shows that the sites that had the most Ascomycota species were detected at Arroyo Seco, Bull Creek, Compton Creek, and Maywood. Maywood had much variability: two points were outliers with high counts >25 , whereas most values were near zero. It is also interesting to note that more than 50 taxa of Ascomycota were identified only to the family level, and some of these may represent heretofore uncharacterized Ascomycetes. Based on these results, an interesting junction of the L.A. River to investigate Ascomycete sequences to a deeper level would be Arroyo Seco and Maywood, which were geographically connected.

The plot of alpha diversity for all fungi, given in Figure 3, shows which sites had the most different types of fungi in any division. Overall, there were 132 taxa of fungi identified. Arroyo Seco, Bull Creek, Compton Creek, Maywood, and Verdugo Wash accumulated the most taxa. An interesting aspect regarding this point is that out of the 132 taxa of fungi, over 30% were Ascomycetes identified only to the family level.

The COI marker performed well in terms of median sequences per sample, which was 18,555. As shown in Table 3, the branch length mean is about 200 units longer for NJ and the variance is about 275,000 units higher for neighbor joining, with respect to the COI marker. For both tree topologies, $k = 3$ is apparent for the number of clusters in terms of COI sequences identified. This seems to reflect that the animal diversity detected by the assay has less breadth than the biodiversity captured by 16S or FITS in this instance.

In Figure S10, the PCA for the COI DNA sequences that were recovered from the L.A. River sediment samples is shown. The first two principal components capture about 33% of the variation in the data. The COI assay captured a picture of lower diversity for the sequences. Samples score low on PC 2 based on the relative abundance of the *Dicrotendipes* species, i.e., non-biting bloodworms [73]. Additionally, low on PC 2 were samples with a high relative abundance of *Eucypris virens*, a cypridid ostracod [74]. The presence of bloodworms is an indicator that other animals are present in the River and is a positive indicator of ecosystem health. The ostracod *E. virens* is sensitive to heavy metal pollution; therefore, the presence of this ostracod in significant numbers is also an indicator of ecosystem health.

The PCA plot for the COI samples color coded by the best PAM clustering is shown in Figure S11. The best PAM clustering in this case was $k = 3$, which exhibited the highest average silhouette width. For the COI sequences, 73 of the samples fall into the first cluster shown in black, ranging from Bowtie Parcel to Verdugo Wash. The second cluster, in red, is composed of Glendale and Sepulveda sediment samples. The third cluster, shown in green, is made up of only two samples from Tujunga Wash and Glendale. This supports the observation that samples were similar to this marker.

The abundance of sequences per taxon for 12S was lower than the other markers assayed at a maximum of only 31,898. Furthermore, the median number of sequences per sample was 953. As shown in Table 3, the branch length means differ for NJ and UPGMA. The UPGMA mean branch length is 1585, whereas the NJ branch length is around 600. The variance is higher for neighbor joining, for the 12S marker, consistent with the other markers. For the NJ tree topology, $k = 2$ appears to be the number of clusters, whereas for UPGMA, $k = 3$ is apparent for the number of clusters in terms of 12S sequences identified.

In Figure S12, the PCA for the 12S DNA sequences that were recovered from the L.A. River sediment samples is given. The first two principal components capture about 63% of the variation in the data. Samples appeared similar in this assay, except for the sample from high on PC 2 in the Elysian Valley that contained a high relative abundance of salmon sequences, which appeared to be an error. In that case, since the taxon is too rare among samples, it could be excluded from the analysis because it might be an error or was unlikely to be relevant to many individuals in the population. Figure S13 shows the PCA plot for the 12S samples color coded by the best PAM clustering, which was $k = 5$, with the highest average silhouette width. A total of 79 out of 90 samples fall into the first cluster, shown in black. The second cluster is mostly made up of Sepulveda Dam sediment DNA samples. The first and third clusters were similar to one another. The fifth cluster, in light blue, is made up of a single sample from Long Beach.

The observed plant alpha diversity for each of the L.A. River sites is plotted in Figure 5. The median number of assigned sequences per sample was relatively low for the plant ITS assay at 9642, although it was not the lowest of all markers. Nevertheless, the number of sequences per taxa had a high maximum at 238,793. As shown in Table 3, the branch length means were similar for NJ and UPGMA, and the variance is about 250,000 units higher for neighbor joining, with respect to the PITS marker. For both tree topologies, $k = 4$ is reflected for the number of clusters in terms of plant sequences identified.

In Figure S14, it is possible to view the PCA for the plant DNA sequences that were recovered from the L.A. River sediment samples. These data are interesting in terms of assessing ecosystem diversity and nitrogen cycling. The first two principal components capture about 34% of the variation in the data. One of the samples from Elysian Valley is high on PC4 due to a high abundance of *Paspalum distictum* sequences. This is a knotgrass found in most of the southern US and the Pacific northwest, where it is native but can become weedy [75]. *Paspalum* plays a role in wetland restoration, since it tolerates waterlogged and saline environments, as well as providing food for deer [75]. Samples from Arroyo Seco are high on PC 3 based on the differential abundance of *Alnus rhombifolia* sequences. Interestingly, most of the *Alnus* sequences were derived from a Tujunga Wash sample. White alders are native to streamside habitats in the US [76]. Alders have been shown to

be key to nitrogen cycling in riparian environments, since they form an association with *Frankia* bacteria. For that reason, they are better at colonizing disturbed habitats [76].

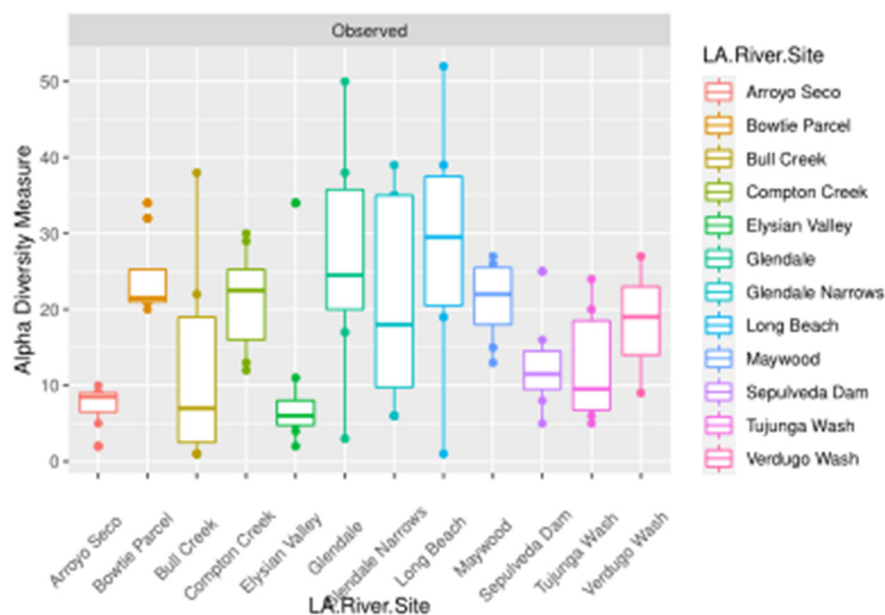


Figure 5. The observed alpha diversity of plant species is depicted in boxplots. This figure answers the question: Which site had the highest number of plant species detected overall? Note that the highest observed alpha diversity tended to be in Glendale, Glendale Narrows, and Long Beach. Again, there is evidence of overdispersion, especially for the Bull Creek, Glendale, and Long Beach samples.

The main factor that separates samples on PC 2 is the abundance of willow species, especially in Bull Creek, Bowtie Parcel, and Arroyo Seco. Most of the *Salix* sequences were derived from two samples from Arroyo Seco. Figure 6 shows the PCA for the plant sequences, color coded by the best PAM clustering. The best PAM clustering for the FITS markers was $k = 4$. The model with four clusters had the highest average silhouette width. The second cluster, shown in red, is composed of Arroyo Seco and Bull Creek. The third cluster consists of sediment samples from Compton Creek-, Sepulveda Dam-, and Glendale-adjacent sites. The fourth cluster, in blue, is made up of Arroyo Seco samples. The first cluster is made up of a mixture of all other samples, which were similar to one another, shown in black.

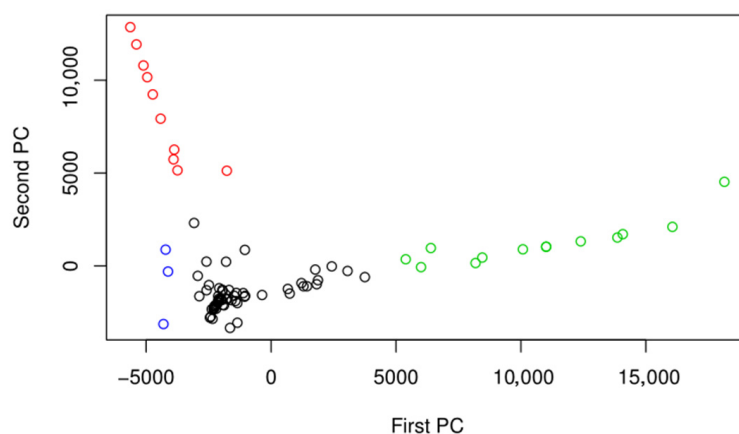


Figure 6. PCA for identified plant sequences from the PITS marker by sample is presented, color coded by the best PAM clustering. Each cluster was color coded according to overall similarity of the plant populations.

4. Discussion

This study has investigated the associations between microorganisms and environmental conditions including soft bottoms versus concrete bottoms, the degree of urbanization, and proximity to a water treatment plant. The physical distance between samples appears to be mirrored by the genetic distance, based on the evidence from PCA with PAM clustering for the 18S markers. Matsuoka et al. found similar results along a river network in Japan in 2019, where they found that fungal DNA assemblages had a spatial structure and samples that were closer to one another tended to be more similar. Overall, our results agree with the numerous studies of urban, eutrophic, and brackish freshwater bodies since proteobacteria, bacteroidetes, firmicutes, cyanobacteria, chloroflexi, actinobacteria, and acidobacteria were all well-represented [77–80]. The elevated presence of Verrucomicrobia and Gammaproteobacteria aligned more with the brackish metagenome [80]. The ostracods detected in high abundance are not known indicator species for heavy metal contamination [81].

In Glendale Narrows, downstream from water reclamation plants, there were abundant cyanobacteria and algae sequences. Eutrophication can lead to hypoxic conditions; since hypoxia can be fatal to fish, this may partly explain the low 12S diversity. The greatest social costs associated with irrigating using reclaimed water are the costs to recreation and the risks to human health due to the potential for the presence of hazardous substances [82]. However, at Glendale Narrows, indicator species for both low nutrient environments and ammonia-abundance were also present. A potential explanation for this is the high abundance of plant species at Glendale Narrows, which assimilate nitrogen. Microbes with nutrient cycling capabilities, such as nitrogen reduction or nitrogen fixation, have been known to be associated with plant growth promotion or may be associated with toxicity. Nevertheless, our results do not agree with Francis et al., 2012, where plant species diversity was expected to decrease in urban environments compared to rural environments [83].

Eukaryotic microbes in the rootzone, such as *Basidiomycota* and *Ascomycota* may help plants with phosphorus solubilization but may be pathogenic to plants or humans. Organisms such as these fungi, which promote phosphorus mineralization, have received less attention over the years [84], although they play important roles in nutrient cycling. Fungi such as *Pleurotus* have been shown to mycoremediate contamination with *E. coli* [85]. The results indicate that L.A. River biome is rich with *Ascomycota* beyond the expected proportion for freshwater bodies, including rivers. *Penicillium* sp. are known to bioaccumulate arsenic and cadmium and are thus mycoremediators of metals [86].

Nitrogen cycling was explained through the differential abundance of ammonia oxidizing archaea; the complete ammonia-oxidizers, *Nitrospira* sp.; nitrate-reducing bacteria, *Marmoricola* sp.; and nitrogen-fixing bacteria, *Devosia* sp., were differentially abundant at soft-bottom sites (p adj < 0.002). The proposed nitrogen cycle for soft-bottom conditions is shown in Figure 7. Ammonia-oxidizing archaea were represented by more than one species. This result partly disagrees with the findings by Cai et al. [87], since ammonia oxidizing archaea were more represented. However, some *Nitrospira* bacteria are complete ammonia oxidizers, so they may be equally important. Interestingly, the results from a recent study indicated that nitrogen pollution in river sediments also contributed to bacteria community shifts [78]. In contrast, the differential abundance of several *cyanobacteria* and other anoxygenic phototrophs was associated with the concrete-bottom sites, which suggested the accumulation of excess nitrogen. *Desulfomicrobium* may play a part in nitrate reduction in concrete environments but conserves more nutrition [66] and is sulfate-dependent [65]. Since denitrification generally requires substrate that is made under aerobic conditions [88], it makes sense that denitrifying bacteria were not as abundant in the concrete environments. Clostridia are indicator species for fecal contamination and sewage [89]. In regard to the reproductive pathogens, as Hervé et al. noted, street gutters are important in the dispersal of putative pathogens from anthropogenic waste [90] and bioremediating species.

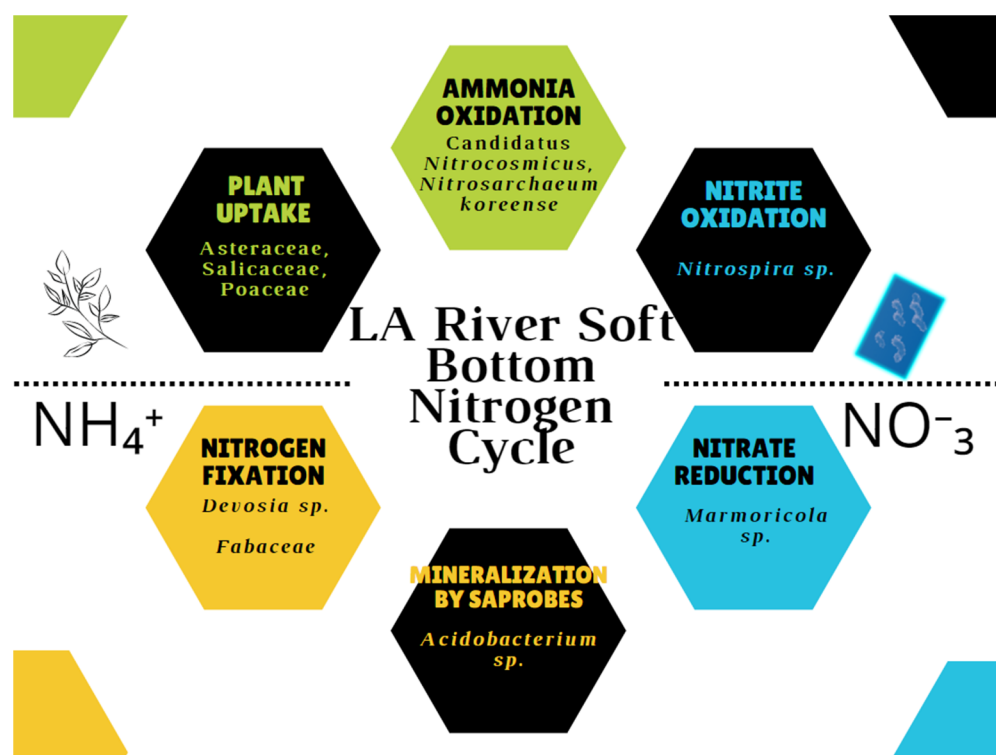


Figure 7. Proposed nitrogen cycle for the L.A. River soft-bottom conditions.

The diversity of cyanobacterial species observed indicated health within the cyanobacteria community. As Stal noted in 2007, *cyanobacteria* are involved in two essential biogeochemical processes on Earth, since they capture both CO_2 and N_2 [88]. *Cyanobacteria* have been known to colonize hostile environments [91] and to produce toxins that bring health risks to the public, such as liver damage, eye irritation, vomiting, and death [92]. However, only 1–2 species of algae were highly represented, which is not an indicator of health for the ecosystem. In a freshwater study by Wang et al., elevated cyanobacteria were associated with bacterioplankton, whereas algae were associated with zooplankton [93]. The heterogeneity and diversity of algae is tied to ecosystem services [94]. According to the Southern California Coastal Water Research Project, *Cladophora* algae support the habitat of wading shorebirds [95]. Treating the underlying anaerobic conditions could promote algal and fish diversity.

The soft-bottom sites tended to be represented by differential abundance of aerobes, whereas the concrete-associated species tended to be alkaliphilic, saliniphilic, calciphilic, sulfate-dependent, and anaerobic. The presence of halophiles is a good indicator of salinity problems. The differential abundance of *Proteobacteria* was associated with soft-bottom sites, and there was an apparent balance in the abundance of organisms responsible for nitrogen cycling.

In recent years, the city of Los Angeles has been reluctant to move toward a soft-bottom channel restoration, since it would necessitate a widening of the channel, which would potentially affect landowners and other infrastructure. Furthermore, although some activists have favored riparian plantings, this also has the potential to slow the flow of water. As the river was channelized in order to decrease flooding risk and efficiently carry away water, the introduction of a vegetative buffer would likely require a widening of the river, and possibly the river's overall footprint. As Levi et al. pointed out, channel restoration benefits appear to be smaller when spread across a larger area [96]. Therefore, this type of effort may be most impactful when applied to the urban stretches that would benefit most from the intervention.

Based on the plant diversity analysis, it was indicated that Maywood had high sequence abundances of weeds such as *Datura*, *Atriplex*, *Oxalis*, and *Chenopodium*, as well as a high abundance of toxic cyanobacteria based on the factor analysis; therefore, Maywood could benefit from the planting of perennial foliage that could also remediate air pollution [97]. According to Liu et al., air pollutants, including particulate matter, nitrous oxide, and carbon monoxide, also influence microbial and fungal communities [98]. Indications tended to suggest that sonicating devices at Maywood and Glendale Narrows for the control of *Cyanobacteria* should be considered, as well as perennial vegetative buffers in Maywood to combat noxious *Datura* plant species and toxic *cyanobacteria* blooms. Interestingly, Maywood samples had differentially abundant *Tetrademus* sp., including *T. obliquus*, which is a phosphorus accumulator and produces valuable lipids for biodiesel [42]. *T. obliquus* may also be used for animal feed; it is known to be rich in amino acids, including the essential amino acid leucine, with a low bioaccumulation of metals [42].

A surprising result is that some sites along the L.A. River were more diverse with plant life than rural Arroyo Seco, especially Bowtie Parcel, Glendale, Long Beach, and Maywood, based on observed alpha diversity. This is most likely due to the landscape plantings of exotic species near Glendale, coastal species at Long beach, and a diverse panel of weed sequences that were identified at Maywood. Plants prevent erosion and create habitats for birds, mammals, invertebrates, amphibians, and reptiles. Plants also help balance nitrogen cycling and can provide a buffer by absorbing some of the nutrients involved in eutrophication. Native plants are useful for bioremediation, soil stabilization, habitat restoration, and as a replacement for invasive species. Native Californian species would also create habitat for wildlife, including birds, insects, and many pollinators. Suggestions for plants for the L.A. River embankment are given in Table 7.

Table 7. Some native Californian plant suggestions for the L.A. River embankment.

Botanical Name	Common Name	Category	Environment
<i>Artemesia douglasiana</i>	Douglas' sagewort	Smaller shrubs and perennials	normal, moist, or saturated soils
<i>Carex praegracilis</i>	field sedge	Smaller shrubs and perennials	normal, moist, or saturated soils
<i>Eleocharis macrostachya</i>	common spikerush	Smaller shrubs and perennials	normal, moist, or saturated soils
<i>Equisetum hyemale</i>	horsetail	Smaller shrubs and perennials	normal, moist, or saturated soils
<i>Juncus patens</i>	common rush	Smaller shrubs and perennials	normal, moist, or saturated soils
<i>Ribes aureum var. gracillimum</i>	golden currant	Smaller shrubs and perennials	normal, moist, or saturated soils
<i>Rosa californica</i>	California wildrose	Smaller shrubs and perennials	normal, moist, or saturated soils
<i>Verbena lasiostachys</i>	vervain	Smaller shrubs and perennials	normal, moist, or saturated soils
<i>Acer negundo</i>	box elder	Larger shrubs and trees	normal, moist, or saturated soils
<i>Acer rhombifolia</i>	white alder	Larger shrubs and trees	normal, moist, or saturated soils
<i>Baccharis salicifolia</i>	mulefat	Larger shrubs and trees	normal, moist, or saturated soils
<i>Juglans californica</i>	black walnut	Larger shrubs and trees	normal, moist, or saturated soils
<i>Platanus racemosa</i>	California sycamore	Larger shrubs and trees	normal, moist, or saturated soils
<i>Populus fremontii</i>	Fremont cottonwood	Larger shrubs and trees	normal, moist, or saturated soils
<i>Salix laevigata</i>	red willow	Larger shrubs and trees	normal, moist, or saturated soils
<i>Salix lasiolepis</i>	arroyo willow	Larger shrubs and trees	normal, moist, or saturated soils
<i>Sambucus mexicana</i>	blue elderberry	Larger shrubs and trees	normal, moist, or saturated soils
<i>Artemesia californica</i>	California sagebrush	Smaller shrubs and perennials	riparian banks, not saturated
<i>Asclepias fasciculata</i>	narrow leaf milkweed	Smaller shrubs and perennials	riparian banks, not saturated
<i>Encelia californica</i>	bush sunflower	Smaller shrubs and perennials	riparian banks, not saturated
<i>Eriogonum fasciculatum</i>	California buckwheat	Smaller shrubs and perennials	riparian banks, not saturated
<i>Lotus scoparius</i>	deerweed	Smaller shrubs and perennials	riparian banks, not saturated
<i>Salvia apiana</i>	white sage	Smaller shrubs and perennials	riparian banks, not saturated
<i>Salvia clevelandii</i>	Cleveland sage	Smaller shrubs and perennials	riparian banks, not saturated
<i>Salvia mellifera</i>	black sage	Smaller shrubs and perennials	riparian banks, not saturated
<i>Baccharis pilularis</i>	coyote brush	Larger shrubs and trees	riparian banks, not saturated

Table 7. Cont.

Botanical Name	Common Name	Category	Environment
<i>Ceanothus</i> spp.	California lilac	Larger shrubs and trees	riparian banks, not saturated
<i>Heteromeles arbutifolia</i>	toyon	Larger shrubs and trees	riparian banks, not saturated
<i>Juglans californica</i>	California walnut	Larger shrubs and trees	riparian banks, not saturated
<i>Manzanita</i> spp.		Larger shrubs and trees	riparian banks, not saturated
<i>Malosma laurina</i>	laurel sumac	Larger shrubs and trees	riparian banks, not saturated
<i>Platanus racemosa</i>	California sycamore	Larger shrubs and trees	riparian banks, not saturated
<i>Rhus integrifolia</i>	lemonade berry	Larger shrubs and trees	riparian banks, not saturated
<i>Sambucus mexicana</i>	blue elderberry	Larger shrubs and trees	riparian banks, not saturated
<i>Quercus agrifolia</i>	coast live oak	Larger shrubs and trees	riparian banks, not saturated

5. Conclusions

Further research should consider the efficacy of sonicating devices at Maywood and Glendale Narrows for the control of *cyanobacteria* [99]. There were poorly characterized microbes and arthropods identified in this study that may present an opportunity for further investigation. These include a possible new species of *Capnioidales* sooty mold in the submerged samples, little known *Chironomidae* lake flies in the Glendale Narrows sample, *Desulfomicrobia* in concrete environments, elusive *Eustigmatophyceae* in Maywood, and unstudied *Verrucomicrobia* and *Flavobacter* in Glendale Narrows. Arroyo Seco and Maywood, which are geographically connected, present an interesting junction of the L.A. River to investigate Ascomycetes and sequence them to a deeper level. This is one of the first attempts to characterize the metagenome of the L.A. River. The diversity and interaction of the bacterial communities with plants and other organisms warrants more attention. The outcomes appear to involve interactions between environmental factors. Further research should consider the functional analysis of similar associations.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/d15070823/s1>.

Author Contributions: Conceptualization, S.S.; methodology, S.S.; formal analysis, S.S.; investigation, S.S., D.M. and R.S.; writing—original draft preparation, S.S.; writing—review and editing, A.E.T., D.M., G.P., S.B., R.S., K.P. and J.F.; visualization, S.S.; supervision, G.P., A.E.T. and S.B. All authors have read and agreed to the published version of the manuscript.

Funding: The authors acknowledge support from the University of California CaleDNA Program who kindly provided us with data prior to publication.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: Research data for the L.A. River Round 1 Project are available from CaleDNA at: https://data.ucedna.com/research_projects/los-angeles-river-round-1/pages/introduction, accessed on 3 April 2023.

Acknowledgments: Thank you to John Creedon and Elias Tarver for technical writing assistance.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. St. Louis Federal Reserve. Federal Reserve Economic Data. Available online: <https://fred.stlouisfed.org/> (accessed on 3 April 2023).
2. United States Census Bureau. Quick Facts: Los Angeles City, California. Available online: <https://www.census.gov/quickfacts/losangelescitycalifornia> (accessed on 3 April 2023).
3. California Environmental Protection Board. Los Angeles River Watershed Impaired Waters. State Water Resources Control Board-Los Angeles. Available online: https://www.waterboards.ca.gov/rwqcb4/water_issues/programs/regional_program/Water_Quality_and_Watersheds/los_angeles_river_watershed/303.shtml (accessed on 6 July 2022).
4. Protecting Our River. Available online: [Protectingourriver.org](https://protectingourriver.org) (accessed on 17 October 2022).
5. Qiu, H.; Gu, L.; Sun, B.; Zhang, J.; Zhang, M.; He, S.; An, S.; Leng, X. Metagenomic Analysis Revealed that the Terrestrial Pollutants Override the Effects of Seasonal Variation on Microbiome in River Sediments. *Bull. Environ. Contam. Toxicol.* **2020**, *105*, 892–898. [CrossRef] [PubMed]

6. Wei, F.; Sakata, K.; Asakura, T.; Date, Y.; Kikuchi, J. Systemic Homeostasis in Metabolome, Ionome, and Microbiome of Wild Yellowfin Goby in Estuarine Ecosystem. *Sci. Rep.* **2018**, *8*, 3478. [CrossRef] [PubMed]
7. Garba, F.; Ogidiaka, E.; Akamagwuna, F.C.; Nwaka, K.H.; Edegbene, O.A. Deteriorating water quality state on the structural assemblage of aquatic insects in a North-Western Nigerian River. *Water Sci.* **2022**, *36*, 22–31. [CrossRef]
8. Yuan, Q.; Wang, P.; Wang, C.; Chen, J.; Wang, X.; Liu, S. Indicator species and co-occurrence pattern of sediment bacterial community in relation to alkaline copper mine drainage contamination. *Ecol. Indic.* **2020**, *120*, 106884. [CrossRef]
9. Khalilova, E.A.; Kotenko, S.T.; Islammagomedova, E.A.; Abakarova, A.A.; Chernyh, N.A.; Aliverdiyeva, D.A. Halophilic bacteria of salt lakes and saline soils of the Peri-Caspian lowland (Republic of Daghestan) and their biotechnological potential. *Vavilov J. Genet. Breed.* **2021**, *25*, 224–233. [CrossRef]
10. Ackerman, D.; Schiff, K.; Trim, H.; Mullin, M. Characterization of water quality in the LA River. *Bull. South. Calif. Acad. Sci.* **2003**, *102*, 17–25.
11. Desfor, G.; Keil, R. Every River Tells a Story: The Don River (Toronto) and the Los Angeles River (Los Angeles) as Articulating Landscapes. *J. Environ. Policy Plan.* **2000**, *2*, 5–23. [CrossRef]
12. Gandy, M. Riparian anomie: Reflections on the Los Angeles River. *Landsc. Res.* **2006**, *31*, 135–145. [CrossRef]
13. Wenger, S.J.; Roy, A.H.; Jackson, C.R.; Bernhardt, E.S.; Carter, T.L.; Filoso, S.; Gibson, C.A.; Hession, W.C.; Kaushal, S.S.; Martí, E.; et al. Twenty-six key research questions in urban stream ecology: An assessment of the state of the science. *J. N. Am. Benthol. Soc.* **2009**, *28*, 1080–1098. [CrossRef]
14. Antwis, R.E.; Griffiths, S.; Harrison, X.; Bou, P.A.; Arce, A.; Bettridge, A.S.; Brailsford, F.; De Menezes, A.; Devaynes, A.; Forbes, K.M.; et al. Fifty important research questions in microbial ecology. *FEMS Microbiol. Ecol.* **2017**, *93*, fix044. [CrossRef]
15. Orsi, J. *Hazardous Metropolis: Flooding and Urban Ecology in Los Angeles*; University of California Press: Berkeley, CA, USA, 2004. Available online: <https://ebookcentral.proquest.com/lib/osu/detail.action?docID=227297> (accessed on 14 February 2023).
16. Halecki, W.; Tomasz, S. Evaluation of Soil Hydrophysical Parameters Along a Semiurban Small River; Soil Ecosystem Services for Enhancing Water Retention in Urban and Suburban Green Areas. *Catena* **2022**, *196*, 104910. [CrossRef]
17. Behera, B.K.; Patra, B.; Chakraborty, H.J.; Sahu, P.; Rout, A.K.; Sarkar, D.J.; Parida, P.K.; Raman, R.K.; Rao, A.R.; Rai, A.; et al. Metagenome analysis from the sediment of river Ganga and Yamuna: In search of beneficial microbiome. *PLoS ONE* **2020**, *15*, e0239594. [CrossRef]
18. Caporaso, J.G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F.D.; Costello, E.K.; Fierer, N.; Gonzalez Peña, A.; Goodrich, J.K.; Gordon, J.I.; et al. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **2010**, *7*, 335–336. [CrossRef]
19. Curd, E.E.; Gold, Z.; Kandlikar, G.S.; Gomer, J.; Ogden, M.; O’Connell, T.; Pipes, L.; Schweizer, T.M.; Rabichow, L.; Lin, M.; et al. Anacapa Toolkit: An environmental DNA toolkit for processing multilocus metabarcode datasets. *Methods Ecol. Evol.* **2019**, *10*, 1469–1475. [CrossRef]
20. Baragaño, B.; Ratié, R.; Sierra, C.; Chrastný, V.; Komárek, K.; Gallego, J. Multiple pollution sources unravelled by environmental forensics techniques and multivariate statistics. *J. Hazard. Mater.* **2021**, *424*, 127413. [CrossRef] [PubMed]
21. Yang, R.H.; Su, J.H.; Shang, J.J.; Wu, Y.Y.; Li, Y.; Bao, D.P.; Yao, Y.J. Evaluation of the ribosomal DNA internal transcribed spacer (ITS), specifically ITS1 and ITS2, for the analysis of fungal diversity by deep sequencing. *PLoS ONE* **2018**, *13*, e0206428. [CrossRef] [PubMed]
22. Caporaso, J.G.; Lauber, C.L.; Walters, W.A.; Berg-Lyons, D.; Huntley, J.; Fierer, N.; Owens, S.M.; Betley, J.; Fraser, L.; Bauer, M.; et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* **2012**, *6*, 1621–1624. [CrossRef] [PubMed]
23. Amaral-Zettler, L.A.; McCliment, E.A.; Ducklow, H.W.; Huse, S.M. A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PLoS ONE* **2009**, *4*, e6372. [CrossRef]
24. Gu, W.; Song, J.; Cao, Y.; Sun, Q.; Yao, H.; Wu, Q.; Chao, J.; Zhou, J.; Xue, W.; Duan, J. Application of the ITS2 Region for Barcoding Medicinal Plants of Selaginellaceae in Pteridophyta. *PLoS ONE* **2013**, *8*, e67818. [CrossRef] [PubMed]
25. Leray, M.; Yang, J.Y.; Meyer, C.P.; Mills, S.C.; Agudelo, N.; Ranwez, V.; Boehm, J.T.; Machida, R.J. A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. *Front. Zool.* **2013**, *10*, 14–34. [CrossRef]
26. White, T.J.; Bruns, T.; Lee, S.; Taylor, J. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In *PCR Protocols: A Guide to Methods and Application*; Innis, M.A., Gelfand, D.H., Sninsk, J.J., White, T.J., Eds.; Academic Press: San Diego, CA, USA, 1990; pp. 315–322.
27. Epp, L.S.; Boessenkool, S.; Bellemain, E.P.; Haile, J.; Esposito, A.; Riaz, T.; Erséus, C.; Gusarov, V.I.; Edwards, M.E.; Johnsen, A.; et al. New environmental metabarcodes for analysing soil DNA: Potential for studying past and present ecosystems. *Mol. Ecol.* **2012**, *21*, 1821–1833. [CrossRef] [PubMed]
28. Borg, I.; Groenen, P. *Modern Multidimensional Scaling. Theory and Applications*; Springer: Berlin/Heidelberg, Germany, 1997.
29. Hastie, T.; Friedman, J. *The Elements of Statistical Learning (ESL)*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2009.
30. McMurdie, P.J.; Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* **2013**, *8*, e61217. [CrossRef] [PubMed]

31. Kandlikar, G.S.; Gold, Z.; Cowen, M.C.; Meyer, R.S.; Freise, A.C.; Kraft, N.J.; Moberg-Parker, J.; Sprague, J.; Kushner, D.J.; Curd, E.E. ranacapa: An R package and Shiny web app to explore environmental DNA data with exploratory statistics and interactive visualizations. *F1000Research* **2018**, *7*, 1734. [CrossRef] [PubMed]
32. Schubert, E.; Rousseeuw, P.J. Fast and eager k-medoids clustering: O (k) runtime improvement of the PAM, CLARA, and CLARANS algorithms. *Inf. Syst.* **2021**, *101*, 101804. [CrossRef]
33. Lepère, C.; Domaizon, I.; Humbert, J.-F.; Jardillier, L.; Hugoni, M.; Debroas, D. Diversity, spatial distribution and activity of fungi in freshwater ecosystems. *PeerJ* **2019**, *7*, e6247. [CrossRef] [PubMed]
34. Anders, S.; Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **2010**, *11*, R106. [CrossRef]
35. Backman, T.W.H.; Girke, T. systemPipeR: NGS workflow and report generation environment. *BMC Bioinform.* **2016**, *17*, 388. [CrossRef]
36. Blighe, K.; Rana, S.; Lewis, M. *EnhancedVolcano: Publication-Ready Volcano Plots with Enhanced Colouring and Labeling*; R Package Version 1.10.0; 2021. Available online: <https://github.com/kevinblighe/EnhancedVolcano> (accessed on 30 March 2023).
37. Maria, L.; Andrés, P. Chapter Eleven-Aminoacyl-tRNA synthetases as drug targets. In *The Enzymes*; de Pouplana, L.R., Kaguni, L.S., Eds.; Academic Press: Cambridge, MA, USA, 2020; Volume 48, pp. 321–350. ISBN 9780128202609. ISSN 1874-6047. [CrossRef]
38. Hanada, S. The Phylum Chloroflexi, the Family Chloroflexaceae, and the Related Phototrophic Families Oscillochloridaceae and Roseiflexaceae. In *The Prokaryotes*; Rosenberg, E., DeLong, E.F., Lory, S., Stackebrandt, E., Thompson, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2014. [CrossRef]
39. Willison, J.C.; Magnin, J.-P. Chapter Eight-Role and Evolution of Endogenous Plasmids in Photosynthetic Bacteria. In *Advances in Botanical Research*; Thomas Beatty, J., Ed.; Academic Press: Cambridge, MA, USA, 2013; Volume 66, pp. 227–265. ISSN 0065-2296; ISBN 9780123979230. [CrossRef]
40. Teneva, I.; Dzhabazov, B.; Koleva, L.; Mladenov, R.; Schirmer, K. Toxic potential of five freshwater Phormidium species (Cyanoprokaryota). *Toxicon* **2005**, *45*, 711–725. [CrossRef]
41. Synechococcus. Microbewiki, Kenyon College. Available online: <https://microbewiki.kenyon.edu/index.php/Synechococcus> (accessed on 1 April 2023).
42. Oliveira, C.Y.B.; Oliveira, C.D.L.; Prasad, R.; Ong, H.C.; Araujo, E.S.; Shabnam, N.; Gálvez, A.O. A multidisciplinary review of *Tetrademus obliquus*: A microalga suitable for large-scale biomass production and emerging environmental applications. *Rev. Aquac.* **2021**, *13*, 1594–1618. [CrossRef]
43. Gouveia, L.; Jazić, J.M.; Ferreira, A.; Maletić, S.; Cvetković, D.; Vidović, S.; Vladi, J.Ć. Green approach for the valorization of microalgae *Tetrademus obliquus*. *Sustain. Chem. Pharm.* **2021**, *24*, 100556. [CrossRef]
44. Montclair State University. Microcystis. New Jersey Center for Water Science and Technology. Available online: <https://www.montclair.edu/water-science/freshwater-cyanobacteria-of-new-jersey/visual-guide-to-cyanobacteria-in-new-jersey/coccoid/colonial/microcystis/> (accessed on 3 July 2022).
45. Chroococcus. Microbewiki, Kenyon College. Available online: https://microbewiki.kenyon.edu/index.php/Chroococcus#Cell_Structure_and_Metabolism (accessed on 14 February 2023).
46. Lucena, T.; Pascual, J.; Garay, E.; Arahall, D.R.; Macián, M.C.; Pujalte, M.J. *Haliea mediterranea* sp. nov., a marine gammaproteobacterium. *Int. J. Syst. Evol. Microbiol.* **2010**, *60*, 1844–1848. [CrossRef] [PubMed]
47. Yang, L.; Tan, Z.; Wang, D.; Xue, L.; Guan, M.-X.; Huang, T.; Li, R. Species identification through mitochondrial rRNA genetic analysis. *Sci. Rep.* **2014**, *4*, 4089. [CrossRef] [PubMed]
48. Kasozi, N.; Kaiser, H.; Wilhelmi, B. Metabarcoding Analysis of Bacterial Communities Associated with Media Grow Bed Zones in an Aquaponic System. *Int. J. Microbiol.* **2020**, *2020*, 8884070. [CrossRef] [PubMed]
49. Islam, T.; Hernández, M.; Gessesse, A.; Murrell, J.C.; Øvreås, L. A Novel Moderately Thermophilic Facultative Methylotroph within the Class Alphaproteobacteria. *Microorganisms* **2021**, *9*, 477. [CrossRef] [PubMed]
50. Jung, M.-Y.; Islam, A.; Gwak, J.-H.; Kim, J.-G.; Rhee, S.-K. *Nitrosarchaeum koreense* gen. nov., sp. nov., an aerobic and mesophilic, ammonia-oxidizing archaeon member of the phylum Thaumarchaeota isolated from agricultural soil. *Int. J. Syst. Evol. Microbiol.* **2018**, *68*, 3084–3095. [CrossRef] [PubMed]
51. Pascual, J.; Foesel, B.; Geppert, A.; Huber, K.J.; Boedeker, C.; Luckner, M.; Wanner, G.; Overmann, J. *Roseisolibacter agri* gen. nov., sp. nov., a novel slow-growing member of the under-represented phylum Gemmatimonadetes. *Int. J. Syst. Evol. Microbiol.* **2018**, *68*, 1028–1036. [CrossRef]
52. Pleurocapsa. Microbewiki, Kenyon College. Available online: [Microbewiki.kenyon.edu/index.php/Pleurocapsa](https://microbewiki.kenyon.edu/index.php/Pleurocapsa) (accessed on 15 March 2023).
53. Hu, Y.; Xing, W.; Song, H.; Zhu, H.; Liu, G.; Hu, Z. Evolutionary Analysis of Unicellular Species in Chlamydomonadales through Chloroplast Genome Comparison with the Colonial Volvocine Algae. *Front. Microbiol.* **2019**, *10*, 1351. [CrossRef]
54. Gorlenko, V.M.; Pierson, B.K. Chloronema. *Bergey's Man. Syst. Archaea Bact.* **2015**, 1–4. [CrossRef]
55. Kurmayer, R.; Christiansen, G.; Holzinger, A.; Rott, E. Single colony genetic analysis of epilithic stream algae of the genus *Chamaesiphon* spp. *Hydrobiologia* **2017**, *811*, 61–75. [CrossRef]
56. Dahal, R.H.; Kim, J. *Altererythrobacter fulvus* sp. nov., a Novel Alkalitolerant Alphaproteobacterium Isolated from Forest Soil. *Int. J. Syst. Evol. Microbiol.* **2018**, *68*, 1502–1508. [CrossRef]
57. Li, H.-P.; Yao, D.; Shao, K.-Z.; Han, Q.-Q.; Gou, J.-Y.; Zhao, Q.; Zhang, J.-L. *Altererythrobacter rhizovicinus* sp. nov., isolated from rhizosphere soil of *Haloxylon ammodendron*. *Int. J. Syst. Evol. Microbiol.* **2020**, *70*, 680–686. [CrossRef]

58. Fidalgo, C.; Rocha, J.; Martins, R.; Proença, D.N.; Morais, P.V.; Henriques, I.; Alves, A. *Altererythrobacter halimionae* sp. nov. and *Altererythrobacter endophyticus* sp. nov., two endophytes from the salt marsh plant *Halimione portulacoides*. *Int. J. Syst. Evol. Microbiol.* **2017**, *67*, 3057–3062. [CrossRef] [PubMed]
59. Gaisin, A.V.; Burganskaya, I.E.; Grouzdev, D.S.; Ashikhmin, A.A.; Kostrikina, N.A.; Bryantseva, I.A.; Koziyeva, V.V.; Gorlenko, V.M. '*Candidatus Viridilinea mediisalina*', a novel phototrophic Chloroflexi bacterium from a Siberian soda lake. *FEMS Microbiol. Lett.* **2019**, *366*, fnz043. [CrossRef]
60. *Calothrix* (Cyanobacteria). Phycokey-Calothrix, University of New Hampshire Center for Freshwater Biology. Available online: cfb.unh.edu/phycokey/Choices/Cyanobacteria/cyano_filaments/cyano_unbranched_fil/tapered_filaments/CALOTHRIX/Calothrix_key.htm#:~:text=Habitat%3A,in%20the%20marine%20intertidal%20zone (accessed on 15 March 2023).
61. Brown, I.; Tringe, S.G.; Ivanova, N.; Goodwin, L.; Shapiro, N.; Alcorta, J.; Pan, D.; Chistoserdov, A.; Sarkisova, S.; Woyke, T.; et al. High-Quality Draft Genome Sequence of the Siderophilic and Thermophilic *Leptolyngbyaceae* Cyanobacterium JSC-12. *Genome Announc.* **2021**, *10*, e0049521. [CrossRef]
62. Fukunaga, Y.; Ichikawa, N. The Class Holophagaceae. In *The Prokaryotes*; Rosenberg, E., DeLong, E.F., Lory, S., Stackebrandt, E., Thompson, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2014. [CrossRef]
63. Senn, S.; Bhattacharyya, S.; Presley, G.; Taylor, A.E.; Stanis, R.; Pangell, K.; Ford, J. Microplastics in the Freshwater Environment. *Ref. Modul. Earth Syst. Environ. Sci.* **2021**, 2023030337. [CrossRef]
64. Yamada, T.; Sekiguchi, Y.; Hanada, S.; Imachi, H.; Ohashi, A.; Harada, H.; Kamagata, Y. *Anaerolinea thermolimosa* sp. nov., *Levilinea saccharolytica* gen. nov., sp. nov. and *Leptolinea tardivitalis* gen. nov., sp. nov., novel filamentous anaerobes, and description of the new classes *Anaerolineae* classis nov. and *Caldilineae* classis nov. in the bacterial phylum *Chloroflexi*. *Int. J. Syst. Evol. Microbiol.* **2006**, *56*, 1331–1340. [CrossRef] [PubMed]
65. Rabus, R.; Venceslau, S.S.; Wöhlbrand, L.; Voordouw, G.; Wall, J.D.; Pereira, I.A. A Post-Genomic View of the Ecophysiology, Catabolism and Biotechnological Relevance of Sulphate-Reducing Prokaryotes. *Adv. Microb. Physiol.* **2015**, *66*, 55–321. [CrossRef] [PubMed]
66. Zhang, T.; Zhuang, X.; Ahmad, S.; Lee, T.; Cao, C.; Ni, S.Q. Investigation of Dissimilatory Nitrate Reduction to Ammonium (DNRA) in Urban River Network Along the Huangpu River, China: Rates, Abundances, and Microbial Communities. *Environ. Sci. Pollut. Res. Int.* **2021**, *29*, 23823–23833. [CrossRef]
67. Keppen, O.I.; Tourova, T.P.; Kuznetsov, B.B.; Ivanovsky, R.N.; Gorlenko, V.M. Proposal of *Oscillochloridaceae* fam. nov. on the basis of a phylogenetic analysis of the filamentous anoxygenic phototrophic bacteria, and emended description of *Oscillochloris* and *Oscillochloris trichoides* in comparison with further new isolates. *Int. J. Syst. Evol. Microbiol.* **2000**, *50*, 1529–1537. [CrossRef]
68. Huber, K.J.; Overmann, J. *Vicinamibacteraceae* fam. nov., the First Described Family within the Subdivision 6 Acidobacteria. *Int. J. Syst. Evol. Microbiol.* **2018**, *68*, 2331–2334. [CrossRef]
69. Zhang, H.; Ma, B.; Huang, T.; Shi, Y. Nitrate reduction by the aerobic denitrifying actinomycete *Streptomyces* sp. XD-11-6-2: Performance, metabolic activity, and micro-polluted water treatment. *Bioresour. Technol.* **2021**, *326*, 124779. [CrossRef]
70. St. Clair, S.; Saraylou, M.; Melendez, D.; Senn, N.; Reitz, S.; Kananipour, D.; Alvarez, A. Analysis of the soil microbiome of a Los Angeles urban farm. *Appl. Environ. Soil Sci.* **2020**, *2020*, 5738237. [CrossRef]
71. WoRMS Editorial Board. World Register of Marine Species. 2021. Available online: <http://www.marinespecies.org> (accessed on 20 April 2021).
72. ITIS Standard Report: Cyprididae. From the Integrated Taxonomic Information System On-Line Database. Available online: <https://www.itis.gov> (accessed on 30 March 2021).
73. Myers, P.; Espinosa, R.; Parr, C.S.; Jones, T.; Hammond, G.S.; Dewey, T.A. The Animal Diversity Web (Online). 2021. Available online: <https://animaldiversity.org> (accessed on 6 January 2021).
74. Smith, R.J.; Martens, K. The ontogeny of the cyprid ostracod *Eucypris virens* (Jurine, 1820) (Crustacea, Ostracoda). In *Evolutionary Biology and Ecology of Ostracoda. Developments in Hydrobiology*; Horne, D.J., Martens, K., Eds.; Springer: Dordrecht, The Netherlands, 2000; Volume 148. [CrossRef]
75. USDA; NRCS. *The PLANTS Database*; National Plant Data Team: Greensboro, NC, USA, 2021. Available online: <http://plants.usda.gov> (accessed on 20 April 2021).
76. White Alder (U.S. National PARK SERVICE). Last Updated 31 January 2020. Available online: <https://www.nps.gov/articles/white-alder.htm> (accessed on 30 March 2023).
77. Gu, Q.; Wang, W.; Li, G.; Liu, Y.; Wong, P.K.; An, T. Pollution profile of waterborne bacterial and fungal community in urban Rivers of Pearl River estuary: Microbial safety assessment. *J. Freshw. Ecol.* **2021**, *36*, 305–322. [CrossRef]
78. Lin, X.; Gao, D.; Lu, K.; Li, X. Bacterial Community Shifts Driven by Nitrogen Pollution in River Sediments of a Highly Urbanized City. *Int. J. Environ. Res. Public Health* **2019**, *16*, 3794. [CrossRef]
79. Wang, P.; Chen, B.; Yuan, R.; Li, C.; Li, Y. Characteristics of aquatic bacterial community and the influencing factors in an urban river. *Sci. Total. Environ.* **2016**, *569–570*, 382–389. [CrossRef]
80. Hugerth, L.W.; Larsson, J.; Alneberg, J.; Lindh, M.V.; Legrand, C.; Pinhassi, J.; Andersson, A.F. Metagenome-Assembled Genomes Uncover a Global Brackish Microbiome. *Genome Biol.* **2015**, *16*, 279. [CrossRef]
81. Iepure, S.; Martinez-Hernandez, V.; Herrera, S.; Rasines-Ladero, R.; de Bustamante, I. Response of microcrustacean communities from the surface—Groundwater interface to water contamination in urban river system of the Jarama basin (central Spain). *Environ. Sci. Pollut. Res.* **2013**, *20*, 5813–5826. [CrossRef] [PubMed]

82. Garcia, X.; Pargament, D. Reusing wastewater to cope with water scarcity: Economic, social and environmental considerations for decision-making. *Resour. Conserv. Recycl.* **2015**, *101*, 154–166. [CrossRef]
83. Francis, R.A. Positioning urban rivers within urban ecology. *Urban Ecosyst.* **2012**, *15*, 285–291. [CrossRef]
84. Li, K.; Hu, J.; Li, T.; Liu, F.; Tao, J.; Liu, J.; Zhang, Z.; Luo, X.; Li, L.; Deng, Y.; et al. Microbial abundance and diversity investigations along rivers: Current knowledge and future directions. *WIREs Water* **2021**, *8*, e1547. [CrossRef]
85. Pini, A.K.; Geddes, P. Fungi Are Capable of Mycoremediation of River Water Contaminated by *E. coli*. *Water Air Soil Pollut.* **2020**, *231*, 2. [CrossRef]
86. Cecchi, G.; Cutroneo, L.; Di Piazza, S.; Vagge, G.; Capello, M.; Zotti, M. From waste to resource: Mycoremediation of contaminated marine sediments in the SEDITERA Project. *J. Soils Sediments* **2019**, *20*, 2653–2663. [CrossRef]
87. Cai, X.; Yao, L.; Wang, S.; Hu, Y.; Dahlgren, R.A.; Wang, Z.; Deng, H.; Liu, L. Properties of ammonia-oxidising bacteria and archaea in a hypereutrophic urban river network. *Freshw. Biol.* **2021**, *67*, 250–262. [CrossRef]
88. Megonigal, J.P.; Hines, M.E.; Visscher, P.T. 8.08-Anaerobic Metabolism: Linkages to Trace Gases and Aerobic Processes. In *Treatise on Geochemistry*; Holland, H.D., Turekian, K.K., Eds.; Elsevier Pergamon: Oxford, UK, 2003; pp. 317–424; ISBN 9780080437514. [CrossRef]
89. Wen, X.; Chen, F.; Lin, Y.; Zhu, H.; Yuan, F.; Kuang, D.; Jia, Z.; Yuan, Z. Microbial Indicators and Their Use for Monitoring Drinking Water Quality—A Review. *Sustainability* **2020**, *12*, 2249. [CrossRef]
90. Hervé, V.; Leroy, B.; Pires, A.D.S.; Lopez, P.J. Aquatic urban ecology at the scale of a capital: Community structure and interactions in street gutters. *ISME J.* **2017**, *12*, 253–266. [CrossRef] [PubMed]
91. Stal, L.J. Cyanobacteria. In *Algae and Cyanobacteria in Extreme Environments. Cellular Origin, Life in Extreme Habitats and Astrobiology*; Seckbach, J., Ed.; Springer: Dordrecht, The Netherlands, 2007; Volume 11. [CrossRef]
92. Bláha, L.; Babica, P.; Maršálek, B. Toxins produced in cyanobacterial water blooms-toxicity and risks. *Interdiscip. Toxicol.* **2009**, *2*, 36–41. [CrossRef]
93. Wang, K.; Razzano, M.; Mou, X. Cyanobacterial blooms alter the relative importance of neutral and selective processes in assembling freshwater bacterioplankton community. *Sci. Total Environ.* **2019**, *706*, 135724. [CrossRef]
94. Caro-Borrero, A.; Carmona-Jiménez, J. The use of macroinvertebrates and algae as indicators of riparian ecosystem services in the Mexican Basin: A morpho-functional approach. *Urban Ecosyst.* **2019**, *22*, 1187–1200. [CrossRef]
95. Assessment of Aquatic Life Use Needs for the Los Angeles River. Los Angeles River Environmental Flows Project, Southern California Coastal Water Research Project. Available online: https://ftp.sccwrp.org/pub/download/DOCUMENTS/TechnicalReports/1154_LARiverAquaticLifeUses.pdf (accessed on 3 April 2023).
96. Levi, P.S.; McIntyre, P.B. Ecosystem responses to channel restoration decline with stream size in urban river networks. *Ecol. Appl.* **2020**, *30*, e02107. [CrossRef]
97. North Dakota State University Extension. Prevent Cyanobacteria Blooms. August 2020. Available online: <https://www.ag.ndsu.edu/news/newsreleases/2020/aug-10-2020/prevent-cyanobacteria-blooms> (accessed on 1 December 2022).
98. Liu, H.; Hu, Z.; Zhou, M.; Hu, J.; Yao, X.; Zhang, H.; Li, Z.; Lou, L.; Xi, C.; Qian, H.; et al. The distribution variance of airborne microorganisms in urban and rural environments. *Environ. Pollut.* **2019**, *247*, 898–906. [CrossRef]
99. US EPA. Control Measures for Cyanobacterial HABs in Surface Water. Available online: <https://www.epa.gov/cyanohabs/control-measures-cyanobacterial-habs-surface-water> (accessed on 30 April 2023).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.