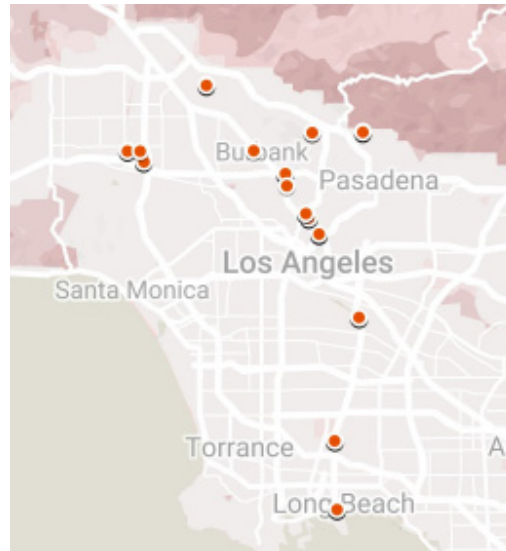Savanah St. Clair Senn

Supplemental Materials

Figure S1. A Google Map of sampling locations is provided, which are marked with a red dot. Map data © 2021 Google.
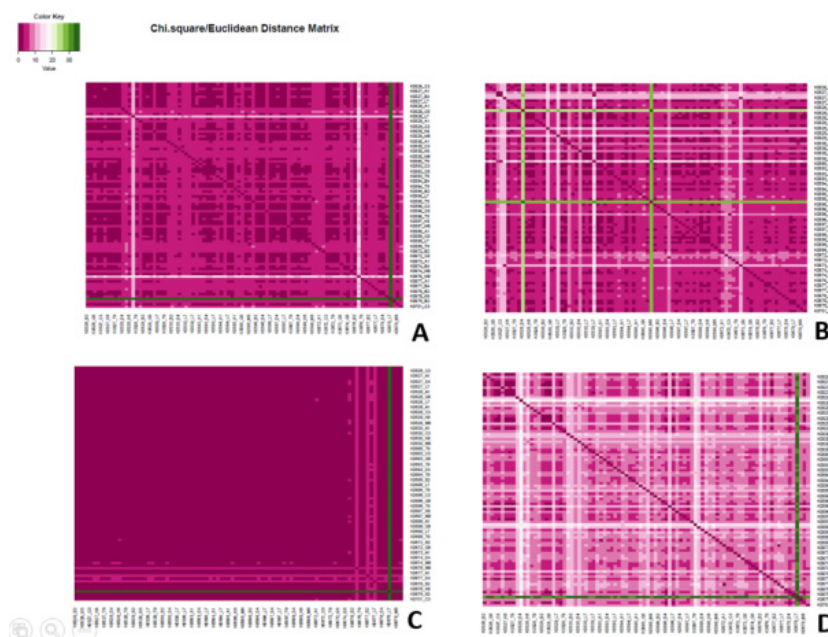

Figure S2. Euclidean Distance Matrix Heatmap visualization of the LA River samples, standardized by the Chi Square Distribution, for 16S (A), 12S (B), CO1 (C), and FITS (D) markers.
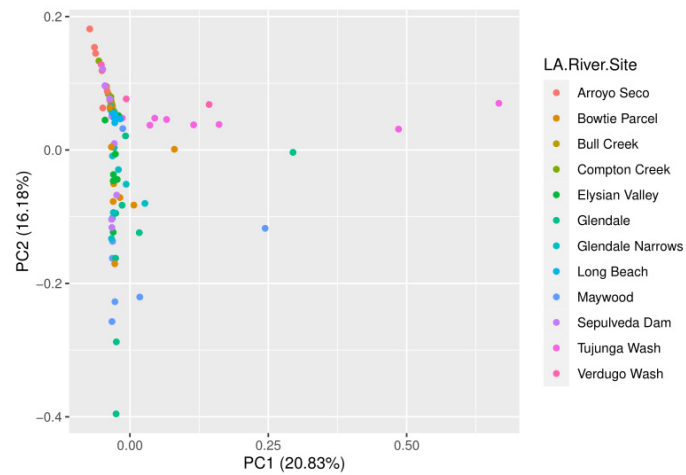
Figure S3. PCA for Fungal identified sequences from the FITS marker by sample, color coded by sampling site. Several samples appear as outliers.
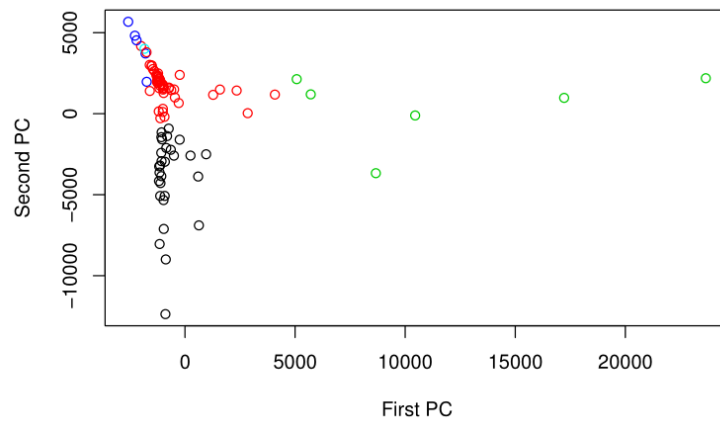

Figure S4. PCA for Fungal identified sequences from the FITS marker is visualized by sample, and color coded by the best PAM clustering.
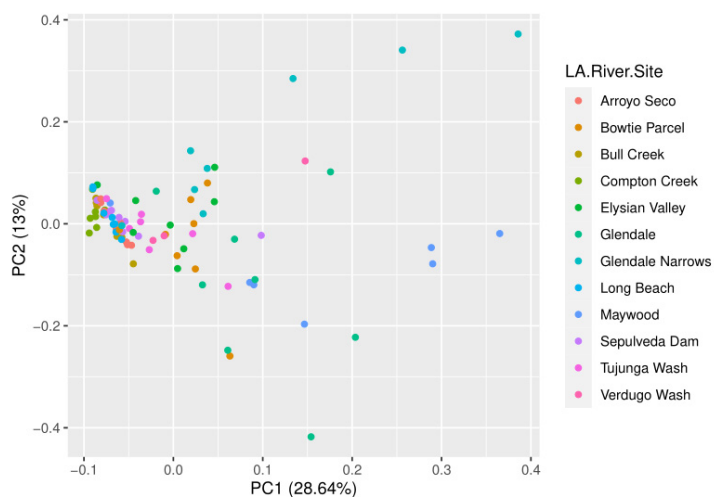

Figure S5. PCA for Bacterial identified sequences from the 16S marker by sample, color coded by sampling site. The evidence of overdispersion was the strongest in the 16S marker count data.
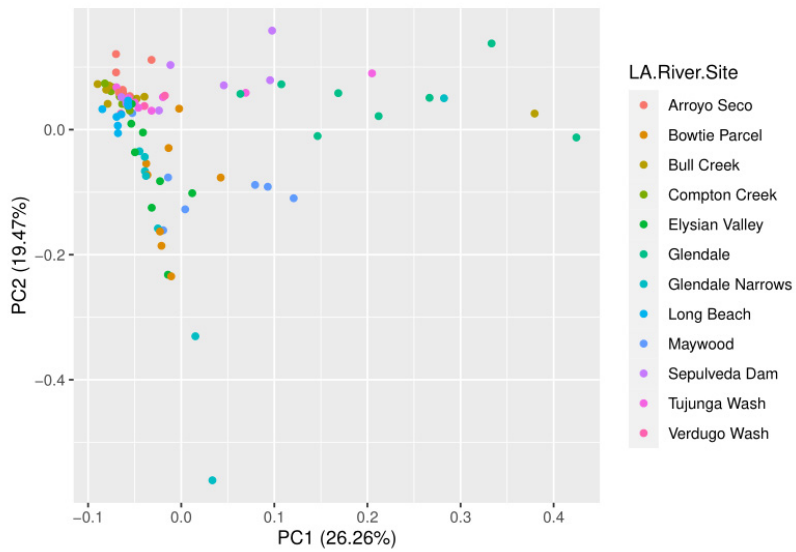
Figure S6. PCA for identified sequences from the 18S marker by sample, color coded by sampling site, which also exhibits overdispersion.
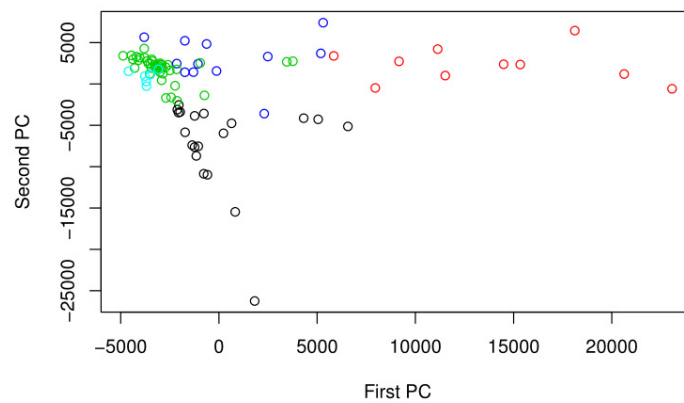

Figure S7. PCA for identified sequences from the 18S marker by sample, color coded by the best PAM clustering.
There is evidence of overdispersion, especially for those samples that are high on both PC 1 and PC 2.
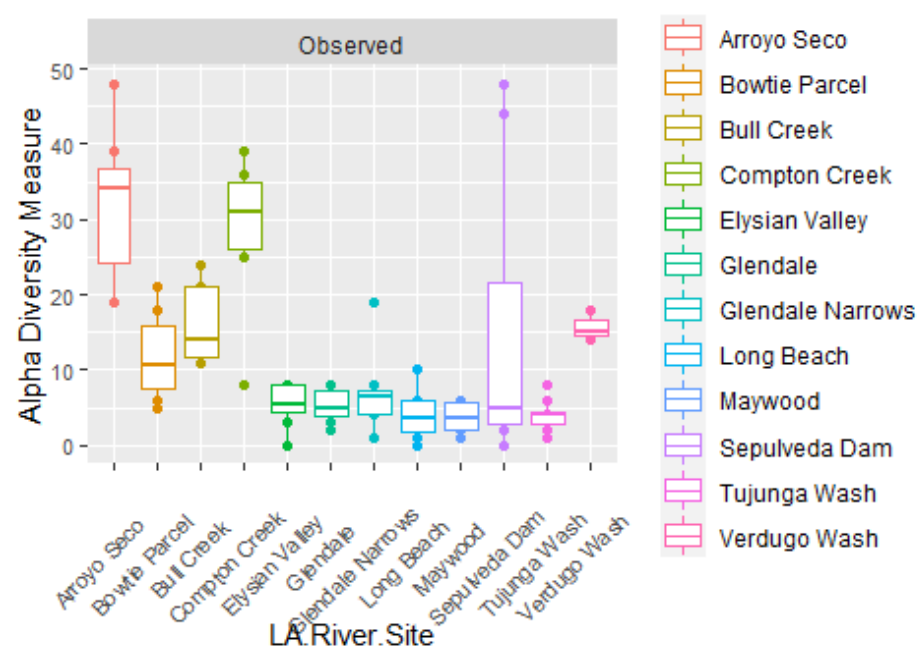
Figure S8. The boxplot of Observed alpha diversity is provided. The figure shows that the species richness for all fungi is the highest for Arroyo Seco, Bull Creek, Compton Creek, Verdugo Wash, and Maywood.
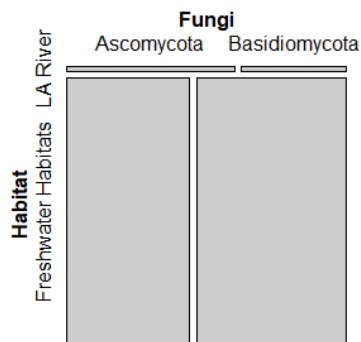

Figure S9. The mosaic plot is given for the Chi square test of proportions. The mosaic plot shows that there is a difference in the proportion of ascomycetes to basidiomycetes in the LA River compared to Freshwater Habitats (Lepere et al 2019).
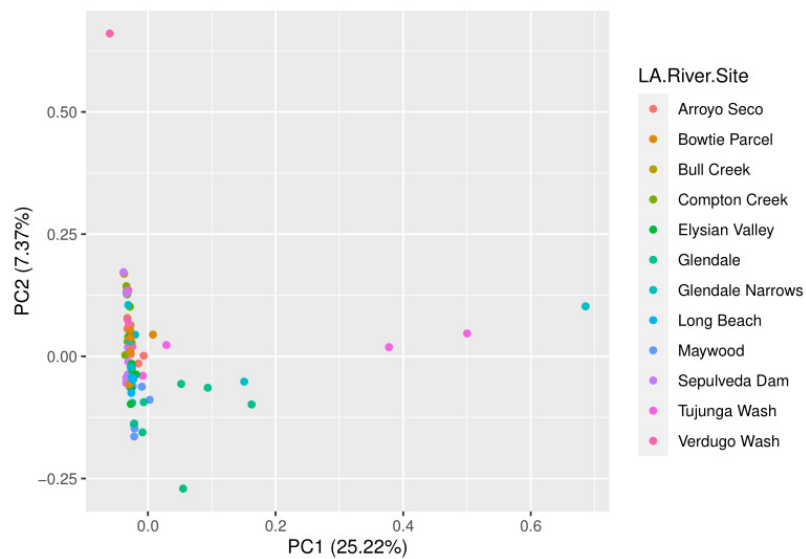
Figure S10. PCA for identified animal sequences from the CO1 marker is visualized by sample, and color coded by sampling site.

Figure S11. PCA for identified animal sequences from the CO1 marker is visualized by sample, and color coded by the best PAM clustering.

Figure S12. PCA for identified sequences from the 12S marker is visualized by sample, and color coded by sampling site.
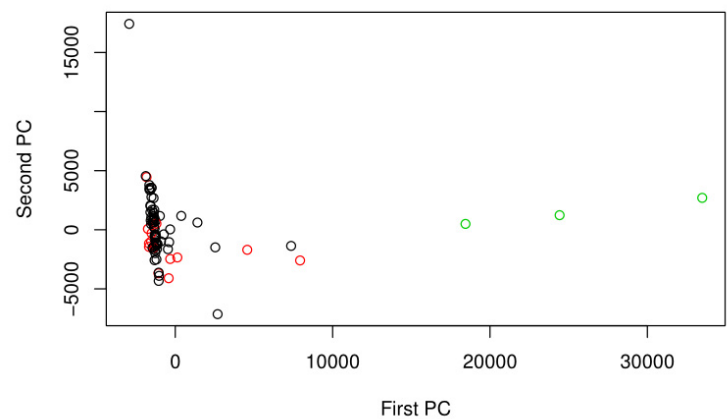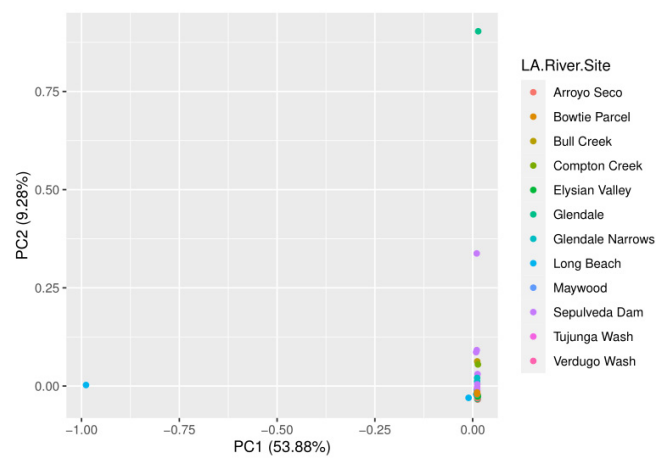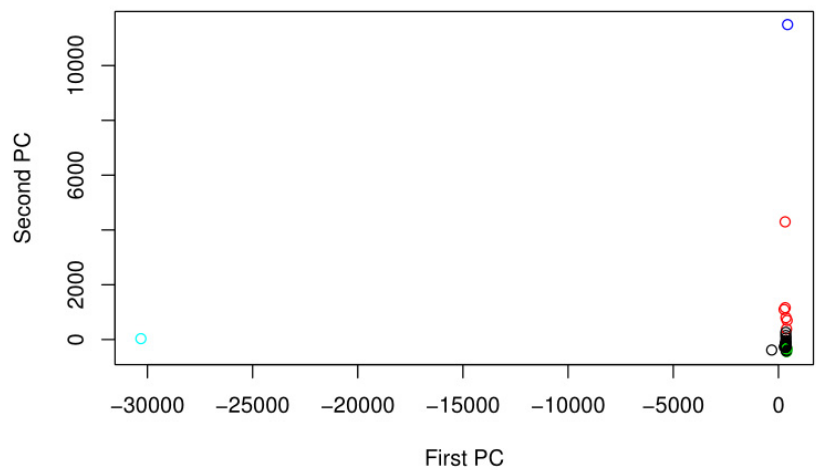

Figure S13. PCA for identified sequences from the 12S marker by is visualized by sample, and color coded by the best PAM clustering. Note that most of the samples appear to be nearly identical.

Figure S14. PCA for identified plant sequences from the PITS marker is visualized by sample, and color coded by sampling site.

Figure S15. Comparison of the relative abundance of Fabaceae and Solanaceae sp. is depicted by sampling site. Which sites contribute to the overall abundance of plant sequences from the *Fabaceae* and *Solanaceae* families? Note that most of the *Fabaceae* sequences arise from Compton Creek. There is also a sizeable proportion of *Fabaceae* sequences from Bowtie Parcel samples. What is the relative abundance of sequences coming from each site? Note the appearance that a small number of sequences of *Datura stramonium* in Arroyo Seco leads to a large abundance of *Datura* sequences in Maywood.

Figure S16. Relative abundance of Alnus, Atriplex, Chenopodium, Paspalum, and Salix sp. color coded by LA River site.

Which sites contribute to the overall abundance of these 5 genera? What is the relative abundance between sites? Notice that *Salix* abundance has contribution from a multitude of samples, whereas *Chenopodium* was mostly found in Maywood and Sepulveda Dam and *Atriplex* was most abundant in Maywood. Elysian Valley, Arroyo Seco, and Tujunga Wash contributed the highest proportions of *Alnus* sequences. *Paspalum* sequences had a high influence on PC3 but apparently were only predominant in Sepulveda Dam, revealing potential bias in the PCA PAM model.

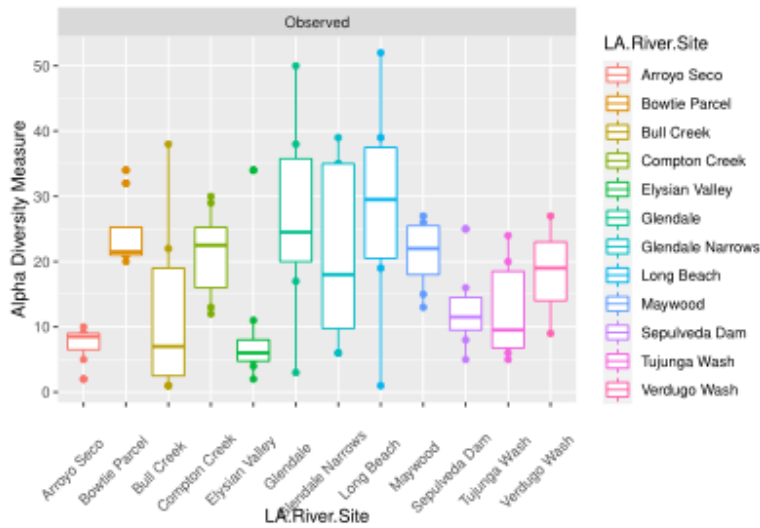Figure S17. The observed alpha diversity of plant species is depicted in boxplots. Which site had the highest number of plant species detected overall? Note that the highest Observed Alpha Diversity tended to be in Glendale, Glendale Narrows, and Long Beach. Again, there is evidence of overdispersion, especially for the Bull Creek, Glendale, and Long Beach samples.
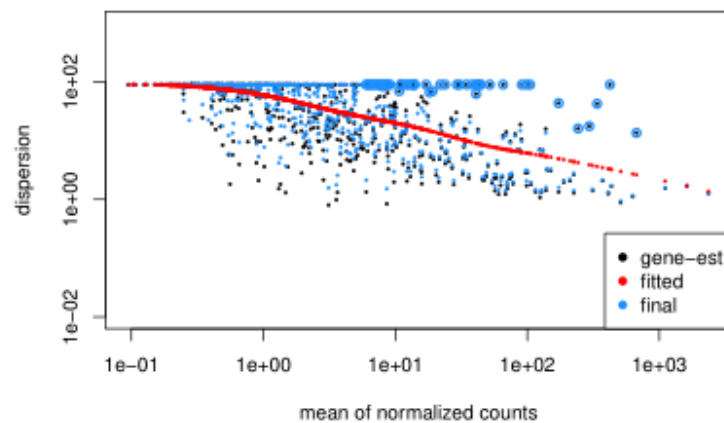

Figure S18. The dispersion estimates for the fungal inter-transcribed spacer (FITS) marker dataset are shown for the local mean method which implements the locfit package. Using the local method, several points are considered as outliers. Notice that those outlying points are not hugged toward the mean in the final estimate, which is represented by the fitted red curve.

Figure S19. The dispersion estimates are shown for the DESeq() method chosen by the algorithm, which appears to be a better fit. Notice that the blue points representing the final estimates are all pulled slightly toward the mean, which is represented by the fitted red curve.

Figure S20. The volcano plot for the The DESeq analysis on the 16S marker data is given. This analysis showed 24 taxa that were differentially abundant between Glendale Narrows and Verdugo Wash, based on significant p-values and Log_2 Fold Changes.

DESeq2 results for FITS

Frequently submerged/intertidal/ marsh vs. Fully Submerged

● NS   ● Log₂ FC   ● p−value and log₂ FC
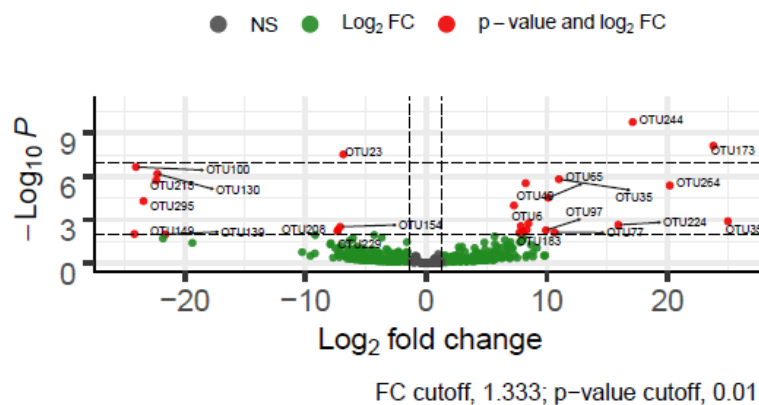

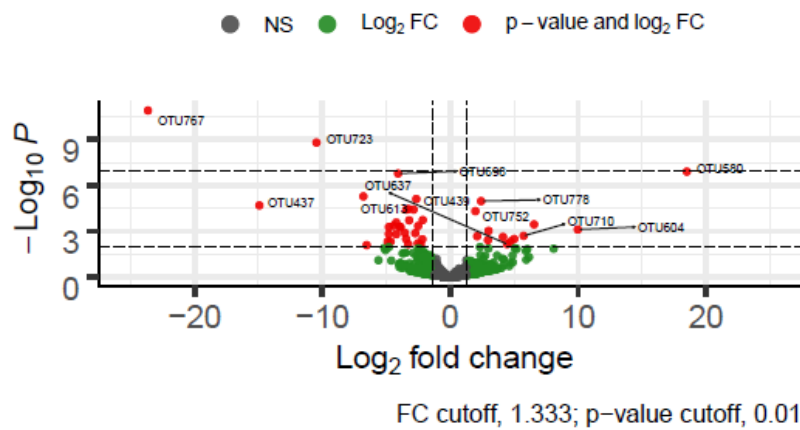FC cutoff, 1.333; p−value cutoff, 0.01

Figure S21. The volcano plot for differential expression of fungi based on the FITS marker is given, with respect to the contrast between frequently submerged versus fully submerged sites.

The differences in fungal taxa as represented by the FITS marker reflected that 27 taxa were differentially expressed, so that some taxa were more likely to be associated with one condition than the other.

DESeq2 results for 16S

Differential abundance: Soft Bottom vs. Concrete

● NS   ● Log₂ FC   ● p−value and log₂ FC

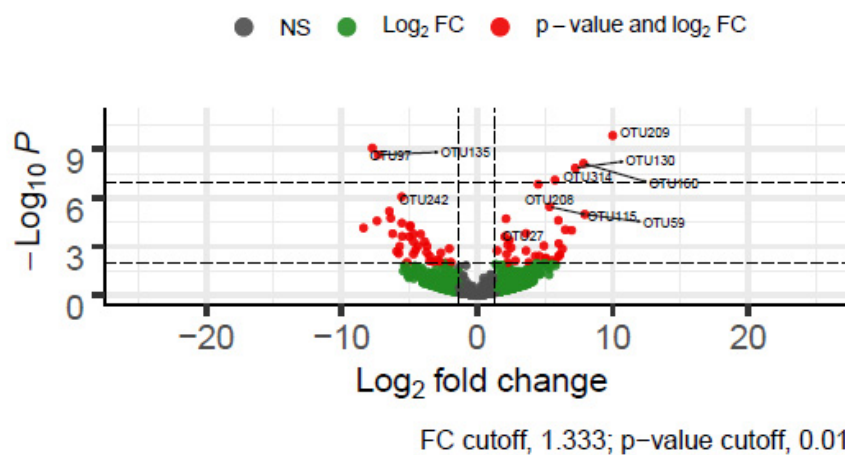FC cutoff, 1.333; p−value cutoff, 0.01

Figure S22. The volcano plot shows a visualization of how many taxa were differentially expressed in the soft bottom and concrete environments, with regards to the bacteria and archaea. The results of the DESeq analysis on the 16S marker seem to reflect two distinct populations associated with soft bottom and concrete environments.

Savanah St. Clair Senn

Table S1. Abundance of ASVs and assigned sequences per sample across the L.A. River sites.

L.A. RIVER	Taxon Abundance			Assigned Seqs/Sample		
Marker	Min	Med	Max	Min	Med	Max
FITS	17	211	183,729	369	18,157	40,447
16S	29	181	109,927	1	15,178	44,190
18S	30	168	299,045	386	24,799	56,966
COI	30	208	153,574	14	18,555	41,257
12S	30	713	31,898	0	953	30,699
PITS	0	265	238,793	133	9642	24,730

Table S2. Status of the outcomes which were investigated further in terms of the specific taxonomic categories that were expected to be found in higher abundance if in fact the hypotheses were supported by strong evidence.

Expectation	Description	Conclusion
Outcome 1a	Differentially higher abundance of fungi in submerged samples	Not supported
Outcome 2a	Differentially higher bacteria abundance in submerged samples	Supported
Outcome 3a	Differentially higher aerobe abundance in soft bottom samples	Partially supported
Outcome 4a	Differentially higher nitrogen reducing species in Glendale Narr.	Not supported
Outcome 5a	Differentially higher ascomycetes in Maywood	Not supported