1   # Agricultural Pollution Risks Influence Microbial Ecology in Honghu Lake

2

3   Maozhen Han[1,2,+], Melissa Dsouza[3, 4, 5,+], Chunyu Zhou[1, +], Hongjun Li[1], Junqian Zhang[6], Chaoyun Chen[1], Qi Yao[1],

4   Chaofang Zhong[1], Hao Zhou[1], Jack A Gilbert[3, 4, 5, *], Zhi Wang[2, *], Kang Ning[1, *]

5

6   *[1] Key Laboratory of Molecular Biophysics of the Ministry of Education, Hubei Key Laboratory of Bioinformatics*

7   *and Molecular-imaging, Department of Bioinformatics and Systems Biology, College of Life Science and*

8   *Technology, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China*

9   *[2] Key Laboratory for Environment and Disaster Monitoring and Evaluation of Hubei, Institute of Geodesy and*

10   *Geophysics, Chinese Academy of Sciences, Wuhan, Hubei 430077, China*

11   *[3] The Microbiome Center, Department of Surgery, University of Chicago, Chicago, IL, 60637, USA*

12   *[4] Argonne National Laboratory, Biosciences Division, Lemont, IL, 60439, USA*

13   *[5] Marine Biological Laboratory, Woods Hole, MA, 02543, USA*

14   *[6] State Key Laboratory of Water Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of*

15   *Sciences, Wuhan, Hubei 430072, China*

16

17

18   [+] These authors contributed equally to this work.

19   [*] Corresponding author. E-mail: ningkang@hust.edu.cn, zwang@whigg.ac.cn, gilbertjack@gmail.com

20

Abstract

Background: Agricultural activities, such as stock-farming, planting industry, and fish aquaculture, can influence the physicochemistry and biology of freshwater lakes. However, the extent to which these agricultural activities, especially those that result in eutrophication and antibiotic pollution, effect water and sediment-associated microbial ecology, remains unclear.

Methods: We performed a geospatial analysis of water and sediment associated microbial community structure, as well as physicochemical parameters and antibiotic pollution, across 18 sites in Honghu lake, which range from impacted to less-impacted by agricultural pollution. Furthermore, the co-occurrence network of water and sediment were built and compared accorded to the agricultural activities.

Results: Physicochemical properties including TN, TP, NO_3^--N, and NO_2^--N were correlated with microbial compositional differences in water samples. Likewise, in sediment samples, Sed-OM and Sed-TN correlated with microbial diversity. Oxytetracycline and tetracycline concentration described the majority of the variance in taxonomic and predicted functional diversity between impacted and less-impacted sites in water and sediment samples, respectively. Finally, the structure of microbial co-associations was influenced by the eutrophication and antibiotic pollution.

Conclusion: These analyses of the composition and structure of water and sediment microbial communities in anthropologically-impacted lakes are imperative for effective environmental pollution monitoring. Likewise, the exploration of the associations between environmental variables (e.g. physicochemical properties, and antibiotics) and community structure is important in the assessment of lake water quality and its ability to sustain agriculture. These results show agricultural practices can negatively influence not only the physicochemical properties, but also the biodiversity

43   of microbial communities associated with the Honghu lake ecosystem. And these results provide

44   compelling evidence that the microbial community can be used as a sentinel of eutrophication and

45   antibiotics pollution risk associated with agricultural activity; and that proper monitoring of this

46   environment is vital to maintain a sustainable environment in Honghu lake.

47   **Keywords:** freshwater, microbial communities, agriculture activities, antibiotics, human impact

48

49   **Background**

50   Water ecosystems, especially inland lakes, have suffered from eutrophication associated with

51   increased agricultural activity comprising fish aquaculture as well as crop and livestock farming on

52   surrounding lands(Brooks et al. 2016, Geist and Hawkins 2016, Williams et al. 2016). Improperly

53   managed agricultural activities, such as excessive and/or improper use of fertilizers and/or pesticides,

54   can cause eutrophication, which can negatively impact biodiversity(Williams et al. 2004). Previous

55   studies have focused on the impact of this pollution on macro-organismal communities(Verdonschot

56   et al. 2011, Williams et al. 2004) and in comparison, microbial ecology remains relatively

57   understudied.

58   Agricultural pollution alters the physicochemical properties of water ecosystems(Baquero et al.

59   2008), which changes the associated microbial community composition and structure. In particular,

60   nitrogen and phosphorus content, water temperature, and pH can fundamentally influence the

61   microbiome(Bowles et al. 2014, Lindström et al. 2005, Xu et al. 2010). However, few studies have

62   quantified the impact of organic pollutants such as herbicides and antibiotics. Determining the

63   ecosystems resilience to such disturbance can aid conservation and help in the development of

64   remediation strategies. There is an urgent need to develop sustainable approaches that establish a

65 balanced relationship between the environment and agricultural production.

66   Antibiotics are widely utilized in livestock and fish aquaculture to promote animal growth and

67 for the prophylactic or curative treatment of infectious disease(Lee et al. 2007), yet surface runoff of

68 the introduction of treated sewage can introduce antibiotic pollution into local water bodies.

69 Antibiotics inhibit microbial activity and can therefore influence biogeochemical processes in these

70 ecosystems(Sengupta et al. 2013) and potentially select for antibiotic resistance mechanisms in

71 environmental bacteria(Cherkasov et al. 2008). In addition, animal sewage can introduce

72 animal-associated antibiotic resistant bacteria into these environments(Wang et al. 2017a), and as

73 such it is necessary to have better quantification of the fitness and recovery rates of these resistant

74 microbes upon release into the environment(Pei et al. 2006).

75   Honghu lake is a large, shallow eutrophic lake with an area of ~350 km^2 and an average depth

76 of ~1.5 m (**Figure 1**). It is located between the irrigation channel of the Four-lake main canal and the

77 Yangzi River. Over the last five decades, Honghu lake has been extensively altered by flood

78 regulation, irrigation, fish aquaculture, shipping, and water supply demands(Ban et al. 2014, Zhang

79 1998). Today, more than 40% of the lake area is used for large-scale aquaculture(Zhang et al. 2017).

80 The intensive use of Honghu lake resources and the emission of sewage and other pollutants

81 including fertilizers, pesticides, and antibiotics into the lake have led to a severe degradation of its

82 water quality and an increase in the frequency of eutrophication events. In 2004, the Honghu Lake

83 Wetland Protection and Restoration Demonstration Project(Zhang et al. 2017) was implemented to

84 ameliorate the negative effects of severe water pollution, and one third of the lake area has been

85 gradually protected under this provision. Consequently, Honghu lake represents a valuable, natural

86 field site for investigating both the efficacy of the restoration program and the long-term effects of

87   agricultural activities, such as the excessive application of antibiotics on water microbial

88   communities.

89   This study aimed to understand the geospatial influence of pollution on the water and

90   sediment-associated microbial communities in Honghu lake. We performed 16S rRNA amplicon

91   sequencing to characterize the microbial ecology, which was correlated against physicochemistry and

92   antibiotic concentrations in these environments. This research was guided by the following scientific

93   questions: (i) How does microbial diversity differ between water and sediment, and how is this

94   influenced by the intensity of agricultural pollution? (ii) Which physicochemical properties and

95   antibiotics are correlated with changes in microbial community structure? (iii) How are the

96   co-occurrence relationships between microbiota influenced by the intensity of agricultural pollution

97   (impacted and less-impacted)? Importantly, this baseline study aims to generate microbioal

98   biomarkers of pollution, and to identify the ecological trends that could be used to provide a sentinel

99   of pollution events in this lake environment.

100

Results

Physicochemical and antibiotic characterization

103   Physicochemical characteristics and antibiotic concentrations were determined for all water and

104   sediment samples (**Figure 1, Supplementary Tables S1, S2, and S4**). Significant differences in pH

105   (t-test, $P=0.0037$), ORP (t-test, $P=0.00068$), TN (t-test, $P=0.035$), and NH_4^+-N (t-test, $P=0.045$)

106   were observed between water samples from impacted and less-impacted (control) sites

107   (**Supplementary Table S1**). Samples from impacted sites were significantly more acidic and had

108   greater concentrations of ORP, TN, and NH_4^+-N when compared to less-impacted sites

109   (**Supplementary Table S1**). Similarly, sediment samples maintained significantly different Sed-OM

110   (*t*-test, *P*=0.002), Sed-LP (*t*-test, *P*=0.0335), Sed-TN (*t*-test, *P*=0.0013), and Sed-TP (*t*-test, *P*=0.021)

111   levels between impacted and less-impacted sites (**Supplementary Table S2**). Less-impacted

112   sediment had greater concentrations of Sed-OM, Sed-LP, and Sed-TN when compared to the

113   impacted sites (**Supplementary Table S2**), which may largely be due to the decomposition of plant

114   material over the preceding winter months. Between impacted and less-impacted sites, the antibiotics

115   ofloxacin (OFL, *t*-test, *P*=0.0079) and sulfamethoxazole (SMZ, *t*-test, *P*=0.043) had significantly

116   different concentrations in water samples, while sulfamerazine (SMR, *t*-test, *P*=0.021) was

117   significantly different in sediment samples (**Supplementary Table S4**); in both cases concentrations

118   were greater in impacted sites.

119

120   *Microbial diversity and community structure*

121   A total of 28 water and sediment samples generated 4,441,405 paired-end 16S rRNA reads,

122   which clustered into 7,785 OTUs (**Supplementary Information**). Microbial alpha diversity was

123   significantly greater in sediment samples (Chao1 (*t*-test, *P*=0.0045, **Supplementary Table S5**) and

124   PD whole tree (*t*-test, *P*=0.003, **Supplementary Table S5**)). The microbial alpha diversity in

125   sediment samples was significantly different between impacted and less-impacted sites (*t*-test,

126   *P*=0.0445, **Supplementary Table S5**). However, no significant difference in alpha diversity was

127   observed in water samples.

128   A total of 53 microbial phyla were identified across all samples (**Figure 2A**), and were

129   differentiated between water and sediment samples (**Figure 2B**), and between impacted and

130   less-impacted sites (PERMANOVA, Bray-Curtis distance, *P* <0.01). In water samples, Proteobacteria

131   (*t*-test, *P* <0.05), Cyanobacteria (*t*-test, *P* <0.05), and Gemmatimonadetes (*t*-test, *P* <0.05) were

132   significantly different between impacted and less-impacted sites (**Figure 2C**). While in sediment

133   samples, Actinobacteria (*t*-test, *P* <0.01), Firmicutes (*t*-test, *P* <0.05), Bacteroidetes (*t*-test, *P* <0.05),

134   Nitrospirae (*t*-test, *P* <0.05), and OP8 (*t*-test, *P* <0.05) were significantly different between impacted

135   and less-impacted sites (**Figure 2D**).

136   Core-OTUs were defined as a set of OTUs that were identified in all samples analyzed, and

137   Pan-OTUs were defined as a set of OTUs that were identified in at least one sample. Core- and

138   Pan-OTUs were determined for all water and sediment samples (**Figure 3**). A total of 132

139   Core-OTUs and 7,418 Pan-OTUs were identified in less-impacted sites, while impacted sites

140   maintained 201 Core-OTUs and 7,706 Pan-OTUs (**Supplementary Figure S3 and Supplementary**

141   **Figure S4**). The Core-OTUs from both the impacted and less-impacted sites were dominated by

142   Proteobacteria, specifically *Janthinobacterium* (**Supplementary dataset 4 and 5**), while

143   Acidobacteria were enriched at the impacted sites (2.79%±1.30%, **Supplementary dataset 4**).

144   Microbial beta diversity was further assessed by UPGMA clustering using the unweighted

145   UniFrac distance matrix. We observed clustering by sampling medium (**Figure 4A** and

146   **Supplementary Figure S5**) and by level of agricultural activity within water and sediment samples

147   (**Figure 4B**). Importantly, greater differences in beta diversity were observed between impacted and

148   less-impacted sites in sediment samples as compared to water samples (**Figure 4B and 4C**).

149

150   ***Functional properties predicted by PICRUSt***

151   We observed clustering of water and sediment microbial communities based on the relative

152   abundance of their predicted functional profiles (**Supplementary Figure S6**) (PERMANOVA,

153   Bray-Curtis distance, $P < 0.0001$). In water samples, functional groups including amino acid related

154   enzymes, peptidases, oxidative phosphorylation, purine metabolism, pyrimidine metabolism, DNA

155   repair and recombination proteins, and arginine and proline metabolism were enriched

156   (**Supplementary Figure S7**). Likwise, in sediment samples we observed an enrichment of functional

157   groups including ribosome biogenesis, secretion system, two-component system, ABC transporters,

158   and pyruvate metaolism (**Supplementary Figure S7**). When investigating agricultural pollution risks,

159   we observed significant differences in the relative abundances of the predicted functional profiles

160   between impacted and less-impacted groups of water samples (PERMANOVA, Bray-Curtis distance,

161    $P < 0.05$). For these samples, the relative abundances of DNA repair and recombination proteins

162   (*t*-test, $P < 0.05$), purine metabolism (*t*-test, $P < 0.05$), secretion systems (*t*-test, $P < 0.05$), oxidative

163   phosphorylation (*t*-test, $P < 0.05$), pyrimidine metabolism (*t*-test, $P < 0.05$), amino acid related

164   enzymes (*t*-test, $P < 0.05$), and arginine and proline metabolism (*t*-test, $P < 0.05$) were significantly

165   different between impacted and less-impacted sites (**Supplementary Figure S7**). In contrast, we

166   observed no significant differences in sediment functional profiles between impacted and

167   less-impacted sites.

168

169   *Correlating physicochemical properties with Microbial diversity*

170   Physicochemical properties including NH_4^+-N, TN, ORP, TP, Tur, COD_{Mn}, and Chl

171   (**Supplementary dataset 6**) were significant explanatory factors that determined the observed

172   clustering pattern of the water microbial communities at impacted sites (**Figure 5A, Supplementary**

173   **Figure S8A**), while pH and DO determined the water microbial community structure at

174   less-impacted sites (**Figure 5A, Supplementary Figure S8A**). For sediment samples, Sed-LP,

175 Sed-TN, and Sed-OM (**Supplementary dataset 7**) were identified as significant explanatory factors

176 shaping the observed clustering pattern at less-impacted sites and Sed-TP for impacted sites (**Figure**

177 **5B, Supplementary Figure S8B**). Based on distance correlations and the statistical significance of

178 Mantel's r-statistic, water physciochemical properties including TN, ORP, NO_3^--N, and NO_2^--N, were

179 strongly correlated with taxonomic and functional composition (**Figure 6A**). For sediment samples,

180 Sed-OM and Sed-TN were strongly correlated with taxonomic data (**Figure 6B**).

181   The antibiotic oxytetracycline (OTC) was the primarily explanatory factor for water microbial

182 diversity variance at impacted sites (**Figure 5C, Supplementary Figure S8A**). While in sediment

183 samples, sulfamerazine (SMR) was the primary factor responsible for the observed clustering of

184 samples, including R1S, R2S, P1S, P3S, and P4S, from impacted sites (**Figure 5D, Supplementary**

185 **Figure S8B**). Mantel correlation assessments were also performed between antibiotic data and

186 compositional data for water and sediment samples (**Figure 6C, Figure 6D,** and **Supplementary**

187 **Figure S9**). The oxytetracycline (OTC) antibiotic class was strongly correlated with water taxonomic

188 and functional composition data (**Figure 6C**), while tetracycline (TC) was strongly correlated with

189 sediment taxonomic and functional composition data (**Figure 6D**). The ofloxacin (OFL) antibiotic

190 class was strongly correlated with taxonomic and functional composition data in sediment samples

191 collected from less-impacted (control) sites (**Supplementary Figure S10B**). Additionally, COD_{Mn}

192 and ciprofloxacin (CIP) were strongly correlated with taxonomic and functional composition in

193 water samples collected from impacted sites (**Supplementary Figure S10C**).

194   Moreover, we observed strong correlations between several OTUs, physicochemical properties,

195 and antibiotic concentrations (**Supplementary Information**). In water samples, *Bacillus flexus*

196 (denovo 71031, **Supplementary dataset 8**) was strongly correlated with TN (r=0.8675,

197   fdr-p=7.89E-5, **Supplementary Figure S11C**), NH$_4^+$-N (r=0.8958, fdr-p=7.89E-5, **Supplementary**

198   **Figure S11D)**, PO$_4^{3-}$-P (r=0.832, fdr-p=2.58E-4, **Supplementary Figure S11E**), and oxytetracycline

199   (OTC, r=0.8381, fdr-p=3.62E-4, **Supplementary Figure S12B**).

200

201   ***Biomarker discovery***

202   In water samples, the LEfSe analysis identified 13 biomarkers for impacted sites and 12 for

203   less-impacted sites. The most differentially abundant bacteria from impacted sites belonged to the

204   phylum Proteobacteria, class Betaproteobacteria and class Gammaproteobacteria (**Figure 7A and**

205   **7B)**. These included members of the orders *Methylophilales*, *Nitrosomonadales*, and *Rhodocyclales*

206   (**Figure 7A and 7**). *Methylophilales* are known for their ability to metabolize methane under aerobic

207   and microaerobic conditions(Beck et al. 2013) and *Nitrosomonadales* are significantly enriched in

208   soils containing high concentrations of N fertilizer(Chávez-Romero et al. 2016). Water samples from

209   less-impacted sites were overrepresented by *Oscillatoriophycideae* and *Synechococcophycideae* in

210   *Cyanobacteria*; and *Saprospiraceae* in *Bacteroidetes* (**Figure 7A and 7B)**.

211   In sediment samples, the LEfSe analysis reported 14 biomarkers enriched in impacted sites and

212   5 enriched in less-impacted sites (**Figure 7C and 7D)**. Biomarkers in samples from impacted sites

213   mainly comprised members of the phylum Actinobacteria, family *Pseudomonadaceae*, order

214   *Burkholderiales*, and *class Flavobacteriia*.. For sediment samples from less-impacted sites, bacteria

215   that were differentially abundant include members of *Paenisporosarcina* genus and candidate family

216   planococcaceae, phylum Firmicutes, order Bacillales, and class Bacilli (**Figure 7 C and 7D)**.

217

218   ***Co-occurrence Network Analysis***

219   Co-occurrence network analysis was performed to visualize and characterize co-occurrence

220   patterns among members of water and sediment microbial communities. The water and sediment

221   network comprised 427 nodes and 189 edges (**Figure 8A**) and 443 nodes and 2,877 edges (**Figure**

222   **8B**), respectively. The density of the water and sediment network was 0.002 and 0.023, respectively.

223   These results suggest that the sediment microbial network was more connected than the water

224   network. Both networks exhibited a scale-free degree distribution pattern, whereby most OTUs had

225   low degree values and fewer hub nodes had high degree values (**Supplementary Figure S15**).

226   We detected modules in water and sediment networks using the WalkTrap community detection

227   algorithm. The modularity of the water and sediment network was 0.878 and 0.559, respectively. A

228   total of 50 clusters with the largest membership of 22 was observed for the water network (**Figure**

229   **8C**). Likewise, for the sediment network we observed a total of 38 clusters with the largest

230   membership of 111 (**Figure 8D**). In the sediment network, most OTUs in module 6 (111 nodes) were

231   members of Anaerolineae of the phylum Chloroflexi and Beta-, Delta-, and Gammaproteobacteria.

232   Additionally, most OTUs in module 7 (81 nodes) of the sediment network were members of the

233   Planococcaceae, a family within the order Bacillales. When compared to the sediment network, we

234   observed fewer and smaller hubs in the water network. In this network, most OTUs in module 2 (22

235   nodes) and module 9 (12 nodes) were members of the genus *Synechococcus* within the order

236   Synechococcales and ACK-M1 within the order Actinomycetales, respectively.

237   We also examined the effect of prolonged agricultural activities on the co-occurrence patterns of

238   water and sediment microbial communities. For this, each node in the water and sediment network

239   was colored as a function of its relative abundance across samples from impacted and less-impacted

240   (control) sites (**Figure 8E and 8F**). In both networks, we observed higher connectedness among

241　OTUs associated with less-impacted samples as compared to those associated with samples from

242　impacted sites. This observation was confirmed when we generated sub-networks for impacted and

243　less-impacted sediment samples by selecting OTUs associated with these samples and all edges

244　among them from the overall sediment co-occurrence network (**Supplementary Figure S16**). We

245　observed higher connectedness in microbes associated with less-impacted samples (measured as

246　node degree, 3.746) as compared to those associated with samples from impacted sites (1.397).

247

Discussions

249　　The extensive application of chemical compounds such as fertilizers, herbicides, and antibiotics,

250　can profoundly influence the cycling and accumulation of nutrients in the sediment and water

251　column of Honghu lake(Chen et al. 2008). These agricultural practices can negatively impact not

252　only the physicochemical properties, but also the biodiversity of microbial communities associated

253　with the lake ecosystem(Baquero et al. 2008). These changes in microbial community composition

254　can in turn affect nutrient cycling and organic matter decomposition, thus impacting overall

255　agricultural productivity.

256　　In our study, we analyzed water and sediment samples from Honghu lake, assessing its

257　microbiome, physicochemical properties, and antibiotic concentrations. We found that despite low

258　human activity, high concentrations of Sed-LP, Sed-TN, and Sed-OM were observed at less-impacted

259　(control) sites, probably due to the abundance of submerged plants. We speculate that the decay of

260　these plants during winter subtantially increases organic matter, total nitrogen(Sand-jensen 1998),

261　and total phosphorus(Horppila and Nurminen 2003) in sediment samples. Hence, as expected from

262　previous research, we found that both water and sediment microbial community structure was

263   correlated with TP and TN concentration(Han et al. 2016, Takamura N 2003). Moreover, in water

264   samples, we observed that *Bacillus flexus* was strongly correlated with TN, NH_4^+-N, PO_4^{3-}-P, and

265   oxytetracycline. More important, previous work on *Bacillus flexus* have shown that members of this

266   species can degrade organic(Guo et al. 2013) and inorganic(Divyashree et al. 2009) nitrogen, thus

267   making it a possible candidate for bioremediation in alkaline wastewater(Wang and Zhao 2013).

268   Some strains of *B. flexus* also demonstrate strong phosphorus solubilization activity(Gechemba et al.

269   2015), and others demonstrated resistance to OTC(Sundaramanickam et al. 2015).

270   As to biomarkers in sediment samples from impacted sites, these included members of the

271   *Hydrogenophaga* genus, belonging to *Burkholderiales* (Class *Betaproteobacteria*), which have been

272   previously associated with agricultural activities(Babujia et al. 2016). Moreover, members of the

273   genus *Pseudomonas*, belonging to family *Pseudomonadaceae*, can play an important role in

274   agricultural ecosystems, particularly those associated with plant growth-promotion and disease

275   suppression were mentioned(Pesaro and Widmer 2006).

276   Co-occurrence network analysis showed that Anaerolineae forms a large component of microbial

277   communities associated with sludge wastewater treatment plants wherein they may play important

278   roles in organic degradation(Nielsen et al. 2009). The phylum Proteobacteria are known to easily

279   metabolize soluble organic substrates(Shao et al. 2013). Among these classes, Deltaproteobacteria, a

280   dominant group often observed in a variety of sediment samples, play an important role in degrading

281   organic compounds to carbon dioxide(Ye et al. 2009). Members of *Synechococcus* are a

282   cosmopolitan cyanobacterium often associated with toxic algal blooms and microcystin

283   production(O'neil et al. 2012, Wood et al. 2017). Likewise, members of ACK-M1, in a recent study,

284   exhibited chemotaxis towards ammonium in a water ecosystem, thus influencing nutrient cycling

285   processes and microbial competitive interactions within this ecosystem(Dennis et al. 2013). The

286   presence of these microbial taxa is indicative of the long-term effect of eutrophication in water

287   environments.

288

Conclusion

290   We analyzed the impacted sites and less-impacted sites of water and sediment samples from

291   Honghu lake and surrounding river and pond sites. The microbiome was analyzed in the context of

292   variable physicochemical properties and antibiotic concentrations. There were significant differences

293   between impacted and less-impacted (control) groups in both water and sediment samples. These

294   differences were observed in physicochemical properties, antibiotic concentration levels, and

295   taxonomic structure. Physicochemical properties including TN, TP, NO3--N, and NO2--N were the

296   main factors driving compositional differences in water samples. Likewise, in sediment samples,

297   Sed-OM and Sed-TN were the main factors driving differences in taxonomic composition. The

298   antibiotics, oxytetracycline and tetracycline were identified as the main drivers of taxonomic and

299   functional structure in water and sediment samples, respectively. As for differences between

300   impacted and less-impacted samples, we identified 25 biomarkers within water communities and 19

301   within sediment microbial communities. Finally, the co-occurrence network analysis revealed

302   differences in co-occurrence patterns by sampling medium (water vs. sediment microbial

303   communities) and by level of agricultural activity (impacted vs. less-impacted microbial

304   communities). These results suggest that continued analyses of the composition and structure of

305   water and sediment microbial communities in such anthropologically-impacted lake environments

306   may provide valuable biomarker data to track pollution. The Honghu Lake Wetland Protection and

307   Restoration Demonstration Project provided preliminary data that highlights the importance of

308   monitoring biodiversity in water micro-ecosystems. Our present work allows further investigation

309   into the impact of agricultural practices on water ecosystems and more importantly, into our ability to

310   remediate these important ecosystems.

311

312   **Materials and Methods**

313   *Sample sites and sampling processes*

314   To investigate the differences in microbial community structure resulting from a wide range of

315   anthropogenic activities, a total of 14 water samples and 14 sediment samples were collected from

316   Honghu lake and surrounding rivers and ponds during 10-11 November 2015. Among these sites, L1

317   is the inlet of inflowing river, and sites L3, L8, L9, and L10 are relatively close to aquaculture areas.

318   Meanwhile, to evaluate the primary source of the antibiotics of Honghu lake, the waters from four

319   major connecting rivers (i.e., R1 to R4) of Honghu lake and four typical aquaculture ponds (i.e., P1

320   to P4), which can exchange water with Honghu lake, were collected(Wang et al. 2017b). In keeping

321   with the Government Protection Zone definition(Zhang et al. 2017) and in taking into account the

322   different sources of pollution at each site(Wang et al. 2017b) (treated sewage, crop, livestock, and

323   fish aquaculture), all sampling sites were categorized into two groups—namely, the impacted and the

324   less-impacted, control group(Zhang et al. 2017). Sampling sites labeled L1, L2, P1, P2, P3, P4, R1,

325   R2, R3, and R4 were classified as impacted, while sites labeled L3, L4, L5, L6, L7, L8, L9, and L10

326   were classified as less-impacted (**Figure 1**).

327   For water sampling, 2 L of water at a depth of 0.3-0.5 m were collected at each site using a

328   cylinder sampler. Approximately 1.5 L of sample was used for physicochemical characterization and

329   antibiotic analysis. The remaining 500 mL of sample was size-fractionated using a 20 μm tulle and a

330   0.22 μm diameter pore size filter membrane (Tianjin Jinteng Experiment Equipment Co., Ltd).

331   Microbial biomass was collected on 0.22 μm diameter pore size filter membranes. These membrane

332   samples were stored onsite in a portable cooler with ice bags, then transported to the laboratory and

333   stored at $-80^{o}C$ until DNA extraction. For sediment sampling, ~200 g of sediment (0-10 cm) was

334   collected at each site and stored in a portable cooler with ice bags until its transportation to the

335   laboratory for subsequent downstream analyses. Approximately 50 g of sediment was used for

336   physicochemical characterization and antibiotic analysis, whilst the remainder was dried in an

337   Ultra-low Freeze Dryer (Christ, German) until no further weight changes were observed. The dried

338   sediment (0.5 g) was used for DNA extraction.

339

340   *Physicochemical characterization and antibiotic analysis*

341   *Physicochemical characterization*: Physicochemical data were collected for all water and

342   sediment samples (**Supplementary Table S1 and Supplementary Table S2**). Physicochemical

343   properties including water temperature (T), pH, dissolved oxygen (DO), temperature compensated

344   conductivity (nlF Cond), salinity (Sal), oxidation-reduction potential (ORP), turbidity (Tur),

345   chlorophyll-*a* (Chl-*a*), and fluorescent dissolved organic matter (fDOM) were measured for all water

346   samples *in situ* by EXO2 (YSI, USA). Additional physicochemical properties including total

347   phosphorus (TP), orthophosphate (PO_4^{3-}-P), total nitrogen (TN), ammonium nitrogen (NH_4^+-N),

348   nitrate nitrogen (NO_3^--N), nitrite nitrogen (NO_2^--N), and potassium permanganate index (oxygen

349   consumption, COD_{Mn}) were assayed as described in previous work (Federation and Association

350   2005). For sediment samples, ORP (Sed-ORP) and pH (Sed-pH) were determined using a pH/ORP

351   portable meter (YSI, USA). Sediment organic matter (Sed-OM) was determined in a muffle furnace

352   at 550°C(Federation and Association 2005). Labile phosphorus (LP), NH_4^+-N (Sed-NH_4^+-N), total

353   phosphorus (Sed-TP), and total nitrogen (Sed-TN) were measured by the NH_4Cl extraction method,

354   the KCl extraction method, the perchloric acid and sulfuric acid digestion method, and the Kjeldahl

355   method, respectively(Bao 2005).

356   *Antibiotic analysis:* Based on a report of antibiotic usage in China(Yang et al. 2016), a total of

357   13 antibiotics were selected for detection in water and sediment samples (**Supplementary Table S4**).

358   These antibiotics can be classified into three groups namely: (i) sulfonamides (SAs), including

359   sulfadiazine (SDZ), sulfamerazine (SMR), sulfamate (SFM), sulfadimidine (SMD),

360   sulfamonomethoxine (SMM), and sulfamethoxazole (SMZ); (ii) fluoroquinolones (FQs), including

361   fleroxacin (FLE), ofloxacin (OFL), ciprofloxacin (CIP), and difloxacin (DIF); and (iii) the

362   tetracycline group (TCs), including tetracycline (TC), oxytetracycline (OTC), and chlortetracycline

363   (CTC). We determined the concentration of these antibiotics in water and sediment samples using a

364   2695 Waters Alliance system (Milford, MA, USA) equipped with an auto sampler-controlled binary

365   gradient system, a micro vacuum degasser, and a 2998 Photodiode Array (PDA) detector. A detailed

366   protocol of the antibiotic extraction process is described in **Supplementary Information**. Of the 13

367   antibiotics that were quantified, nine antibiotics including TC, OTC, CTC, SDZ, SMR, SMD, OFL,

368   CIP, and SMZ were selected for further analysis in this study.

369

370   ***DNA extraction and 16S rRNA gene sequencing***

371   DNA was extracted from water filter membranes and dried sediment using a modified

372   hexadecyltrimethylammonium bromide (CTAB) method(Cheng et al. 2014a, Cheng et al. 2014b,

373   Porebski et al. 1997) (**Supplementary Information**). All extracted DNA was dissolved in TE buffer

374   and stored at −20°C until further use.

375   DNA samples were quantified using a Qubit® 2.0 Fluorometer (Invitrogen, Carlsbad, CA) and

376   DNA quality was assessed on 0.8% agarose gels. Approximately 5-50 ng of DNA was used as

377   template for amplifying the V4-V5 hypervariable region of the 16S rRNA gene of microbiota for

378   each sample. Sequences for the forward and reverse primers are "GTGYCAGCMGCCGCGGTAA"

379   and "CTTGTGCGGKCCCCCGYCAATTC", respectively(Han et al. 2016). The sequencing library

380   was constructed using a MetaVx™ Library Preparation kit (GENEWIZ, Inc., South Plainfield, and

381   NJ, USA). Indexed adapters were added to the ends of the 16S rDNA amplicons by limited cycle

382   PCR. DNA libraries were verified using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo

383   Alto, CA, USA) and quantified by Qubit® 2.0 and quantitative PCR (Applied Biosystems, Carlsbad,

384   CA, USA). All sequencing reactions were performed on the Illumina MiSeq platform using

385   paired-end sequencing technology (2*300 bp).

386

387   *Quality control, OTU clustering, and taxonomy assignment*

388   All 16S rRNA gene amplicons were processed according to the ensuing criteria and sequences

389   below the set quality threshold were excluded from subsequent analyses. Firstly, paired-end reads

390   were spliced using the 'make.contigs' command in mothur(Schloss et al. 2009) (version 1.25.0) with

391   default settings. All reads containing ambiguous base calls (N), those longer than 500 bp, and those

392   shorter than 300 bp were removed. Putative chimeras were identified using the SILVA

393   database(Quast et al. 2013) (Release 123) and removed using the 'chimera.uchime' and

394   'remove.seqs' commands in mothur. All high-quality sequences were aligned using PyNAST and

395   dereplicated with UCLUST (Caporaso et al. 2010a) in QIIME (Quantitative Insights Into Microbial

396   Ecology, v1.9.1)(Caporaso et al. 2010b). Finally, the Greengenes database (version 13_8)(DeSantis

397   et al. 2006) was used as the reference database for classifying *de novo* operational taxonomic units

398   (OTUs) that were clustered at the 97% nucleotide identity threshold. We set 0.001% as the threshold

399   to filter the low-abundance OTUs and keep abundant OTUs, i.e., only OTUs with the read

400   counts >0.001% of the total reads of all samples were kept for analysis(Li et al. 2015).

401

402   *Microbial diversity assessment*

403   Microbial alpha- and beta-diversity values were determined using the QIIME(Caporaso et al.

404   2010b) pipeline. For alpha-diversity, rarefaction curves were drawn based on the following metrics:

405   Observed OTUs, Chao1, Phylogenetic Diversity (PD) Whole Tree metric, and the Shannon evenness

406   metric(Magurran 2013). For beta-diversity analysis, the final OTU table was rarefied to contain

407   61,088 reads per sample. Bray-Curtis, weighted and unweighted UniFrac distance metrics(Lozupone

408   and Knight 2005) were used to measure community similarity between samples. Microbial

409   community clustering was arrayed by Principle Coordinate Analysis (PCoA) and visualized using

410   Emperor(Vázquez-Baeza et al. 2013) in QIIME. The hierarchical clustering method, UPGMA

411   (Unweighted Pair Group Method with Arithmetic Mean), was applied to cluster all water and

412   sediment samples, and the clustering tree was visualized in FigTree (version 1.4.2)(Rambaut 2014).

413   Permutational multivariate analysis of variance (PERMANOVA)(Anderson 2001) was performed on

414   the Bray-Curtis distance matrix to compare differences in community structure.

415

416   *Functional profiling*

417   PICRUSt (version 1.0.0-dev)(Langille et al. 2013) was used to make functional predictions

418   based on the 16S rDNA dataset from each sample. For this, OTU-picking was performed on all

419   quality-filtered sequence data using the 'pick_closed_reference_otus.py' command in QIIME. OTUs

420   were clustered at the 97% nucleotide identity threshold using the Greengenes database. The OTU

421   table was normalized using the 'normalize_by_copy_number.py' command. The normalized OTU

422   table was used for functional prediction with the 'predict_metagenomes.py' script, and functional

423   trait abundances were determined for each sample using the KEGG database (version 66.1, May 1,

424   2013)(Kang et al. 2016). Finally, the predicted functional content was collapsed to level three of the

425   KEGG hierarchy using the 'categorize_by_function.py' script.

426

Analysis of the relationships between physicochemical properties, antibiotics, and microbial communities

429   Canonical correspondence analysis (CCA) was used to identify an environmental basis for

430   community ordination, revealing relationships between microbial communities and environmental

431   factors(Ter Braak 1986). For this, the CCA function in R package, vegan was utilized. We utilized

432   the 'envfit' function(Dawson et al. 2012, Virtanen et al. 2009) with 999 permutations to reveal

433   significant correlations between physicochemical properties, antibiotics, and microbial communities.

434   To further investigate correlations between environmental factors (including physicochemical

435   properties and antibiotics) and OTUs, we applied a low-abundance filter to remove OTUs whose

436   relative abundance did not exceed 0.01% in any sample (as previously reported by (Sunagawa et al.

437   2015)). Similarly, for physicochemical data and antibiotics data, the values of each variable were

438   transformed to z-scores(Crocker and Algina 1986), based on which the Pearson Correlation

439   Coeffcient between each environmental factor and each OTU was calculated. To select for significant

440   interactions between an environmental factor and an OTU, the threshold of the r-value and the False

441   Discovery Rate (FDR)-corrected-*p*-value of the Pearson Correlation Coeffcient was set at 0.8 and

442   0.05, respectively.

443

Analysis of environmental drivers of microbial community composition

445   We noted environmental drivers of microbial community composition on the basis of (i)

446   compositional data, which includes taxonomic composition (relative taxonomic abundances) and

447   functional composition at KEGG module level three; (ii) physicochemical data; and (iii) antibiotics

448   data. To pre-process compositional data, we applied a low-abundance filter to remove OTUs whose

449   relative abundance did not exceed 0.01% in any sample and then log transformed the relative

450   abundances. Likewise, for physicochemical and antibiotics data, the values of each variable were

451   transformed to *z*-scores. Based on the Euclidean distances, we computed Mantel correlations

452   between the physicochemical data and compositional data and then the antibiotics data and

453   compositional data (9,999 permutations). We obtained the results in R (version 3.3.1) and visualized

454   it in the Adobe Illustrator (version 16.0.0). Taxonomic composition and functional composition data

455   were correlated to each antibiotic and physicochemical property by Mantel's tests. The distance

456   correlations and the statistical significance of Mantel's r statistic corresponded to edge width and

457   edge color, respectively(Sunagawa et al. 2015).

458

Biomarker analysis

460   Based on their location, all water and sediment samples can be divided into two

461   groups—impacted and less-impacted (control) groups. It is well known that the taxonomic

462   composition of a microbial community can be impacted by local environmental variables. As a result,

463   some bacteria might be enriched by distinctive environmental states. Linear discriminate analysis

464   (LDA) effect size (LEfSe)(Segata et al. 2011) was used to select biomarkers in impacted and

465   less-impacted (control) groups in water and sediment samples. Briefly, the taxa abundance table was

466   imported into the LEfSe pipeline, and the parameters were set as follows: the alpha value for the

467   factorial Kruskal-Wallis test(Breslow 1970) among classes and the p-value for the pairwise Wilcoxon

468   test between subclasses were both chosen to be 0.05. The threshold for the logarithmic LDA score

469   for discriminative features was set at 3.0 and 3.5 for water and sediment samples, respectively.

470

471   *Co-occurrence network analysis*

472   To reduce sparsity, we selected water and sediment OTUs that were present in at least 50% of

473   all water and sediment samples, respectively. We then generated separate networks for water and

474   sediment microbial communities. The co-occurrence network was constructed using the CAVNet

475   package(Cardona 2017) in R (as previously described by(Ma et al. 2016)). Briefly, water and

476   sediment networks were inferred using the Spearman correlation matrix with the WGCNA

477   package(Langfelder and Horvath 2012). In this network, co-occurring OTUs are represented by

478   nodes and connected by edges. The network deconvolution method was utilized to distinguish direct

479   correlation dependencies(Feizi et al. 2013). All p-values were corrected for multiple testing using the

480   Benjamini and Hochberg FDR-controlling procedure(Benjamini et al. 2006). The cutoff of the

481   FDR-corrected-p-value was set at 0.01. Random matrix theory-based methods were utilized to

482   determine the cutoff of Spearman correlation coefficients for water (0.84) and sediment (0.81)

483   networks. All network properties were calculated using the igraph package in R(Csardi and Nepusz

484 2006). We also utilized igraph to visualize and generate water and sediment networks. The WalkTrap

485 community detection algorithm was used to identify modules in water and sediment networks(Pons

486 and Latapy 2005). To study the effect of prolonged agricultural practices, we colored each node

487 within the water and sediment network as function of its relative abundance at impacted and

488 less-impacted (control) sites using the 'plot_network_by_continuous_variable' function in CAVNet.

489

References

491 Anderson, M.J. (2001) A new method for non-parametric multivariate analysis of variance. Austral Ecology 26(1),
492 32-46.

493 Babujia, L.C., Silva, A.P., Nakatani, A.S., Cantão, M.E., Vasconcelos, A.T.R., Visentainer, J.V. and Hungria, M.
494 (2016) Impact of long-term cropping of glyphosate-resistant transgenic soybean [Glycine max (L.) Merr.] on soil
495 microbiome. Transgenic Research, 1-16.

496 Ban, X., Wu, Q., Pan, B., Du, Y. and Feng, Q. (2014) Application of Composite Water Quality Identification Index
497 on the water quality evaluation in spatial and temporal variations: a case study in Honghu Lake, China.
498 Environmental monitoring and assessment 186(7), 4237-4247.

499 Bao, S. (2005) Agricultural and chemistry analysis of soil. Agric. Press, Beijing, China.

500 Baquero, F., Martínez, J.-L. and Cantón, R. (2008) Antibiotics and antibiotic resistance in water environments.
501 Current Opinion In Biotechnology 19(3), 260-265.

502 Beck, D.A., Kalyuzhnaya, M.G., Malfatti, S., Tringe, S.G., del Rio, T.G., Ivanova, N., Lidstrom, M.E. and
503 Chistoserdova, L. (2013) A metagenomic insight into freshwater methane-utilizing communities and evidence for
504 cooperation between the Methylococcaceae and the Methylophilaceae. PeerJ 1, e23.

505 Benjamini, Y., Krieger, A.M. and Yekutieli, D. (2006) Adaptive linear step-up procedures that control the false
506 discovery rate. Biometrika, 491-507.

507 Bowles, T.M., Acosta-Martínez, V., Calderón, F. and Jackson, L.E. (2014) Soil enzyme activities, microbial
508 communities, and carbon and nitrogen availability in organic agroecosystems across an intensively-managed
509 agricultural landscape. Soil Biology and Biochemistry 68, 252-262.

510 Breslow, N. (1970) A generalized Kruskal-Wallis test for comparing K samples subject to unequal patterns of
511 censorship. Biometrika 57(3), 579-594.

512 Brooks, B.W., Lazorchak, J.M., Howard, M.D., Johnson, M.V.V., Morton, S.L., Perkins, D.A., Reavie, E.D., Scott,
513 G.I., Smith, S.A. and Steevens, J.A. (2016) Are harmful algal blooms becoming the greatest inland water quality
514 threat to public health and aquatic ecosystems? Environmental Toxicology and Chemistry 35(1), 6-13.

515 Caporaso, J.G., Bittinger, K., Bushman, F.D., DeSantis, T.Z., Andersen, G.L. and Knight, R. (2010a) PyNAST: a
516 flexible tool for aligning sequences to a template alignment. Bioinformatics 26(2), 266-267.

517 Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G.,
518 Goodrich, J.K. and Gordon, J.I. (2010b) QIIME allows analysis of high-throughput community sequencing data.
519 Nature Methods 7(5), 335-336.

520 Cardona, C. (2017) CAVNet.

521 Chávez-Romero, Y., Navarro-Noya, Y.E., Reynoso-Martínez, S.C., Sarria-Guzmán, Y., Govaerts, B., Verhulst, N.,
522 Dendooven, L. and Luna-Guido, M. (2016) 16S metagenomics reveals changes in the soil bacterial community
523 driven by soil organic C, N-fertilizer and tillage-crop residue management. Soil and Tillage Research 159, 1-8.
524 Chen, M., Chen, J. and Sun, F. (2008) Agricultural phosphorus flow and its environmental impacts in China.
525 Science of The Total Environment 405(1), 140-152.
526 Cheng, X., Chen, X., Su, X., Zhao, H., Han, M., Bo, C., Xu, J., Bai, H. and Ning, K. (2014a) DNA extraction
527 protocol for biological ingredient analysis of LiuWei DiHuang Wan. Genomics, Proteomics & Bioinformatics
528 12(3), 137-143.
529 Cheng, X., Su, X., Chen, X., Zhao, H., Bo, C., Xu, J., Bai, H. and Ning, K. (2014b) Biological ingredient analysis
530 of traditional Chinese medicine preparation based on high-throughput sequencing: the story for Liuwei Dihuang
531 Wan. Scientific Reports 4.
532 Cherkasov, A., Hilpert, K., Jenssen, H., Fjell, C.D., Waldbrook, M., Mullaly, S.C., Volkmer, R. and Hancock, R.E.
533 (2008) Use of artificial intelligence in the design of small peptide antibiotics effective against a broad spectrum of
534 highly antibiotic-resistant superbugs. ACS chemical biology 4(1), 65-74.
535 Crocker, L. and Algina, J. (1986) Introduction to classical and modern test theory, ERIC.
536 Csardi, G. and Nepusz, T. (2006) The igraph software package for complex network research. InterJournal,
537 Complex Systems 1695(5), 1-9.
538 Dawson, K.S., Strąpoć, D., Huizinga, B., Lidstrom, U., Ashby, M. and Macalady, J.L. (2012) Quantitative
539 fluorescence in situ hybridization analysis of microbial consortia from a biogenic gas field in Alaska's Cook Inlet
540 Basin. Applied and Environmental Microbiology 78(10), 3599-3605.
541 Dennis, P.G., Seymour, J., Kumbun, K. and Tyson, G.W. (2013) Diverse populations of lake water bacteria exhibit
542 chemotaxis towards inorganic nutrients. The ISME Journal 7(8), 1661-1664.
543 DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P. and
544 Andersen, G.L. (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with
545 ARB. Applied and Environmental Microbiology. 72(7), 5069-5072.
546 Divyashree, M.S., Rastogi, N.K. and Shamala, T.R. (2009) A simple kinetic model for growth and biosynthesis of
547 polyhydroxyalkanoate in Bacillus flexus. New Biotechnology 26(1), 92-98.
548 Federation, W.E. and Association, A.P.H. (2005) Standard methods for the examination of water and wastewater.
549 American Public Health Association (APHA): Washington, DC, USA.
550 Feizi, S., Marbach, D., Médard, M. and Kellis, M. (2013) Network deconvolution as a general method to
551 distinguish direct dependencies in networks. Nature biotechnology 31(8), 726-733.
552 Gechemba, O.R., Budambula, N., Makonde, H.M., Julius, M. and Matiru, V.N. (2015) Potentially beneficial
553 rhizobacteria associated with banana plants in Juja, Kenya. Journal of Biodiversity and Environmental Sciences
554 7(2), 181-188.
555 Geist, J. and Hawkins, S.J. (2016) Habitat recovery and restoration in aquatic ecosystems: current progress and
556 future challenges. Aquatic Conservation: Marine and Freshwater Ecosystems.
557 Guo, D., Guan, L., Zhang, C., Wang, X. and Shan, L. (2013) UV Mutagenesis Breeding of Bacillus flexus Highly
558 Degrading Organic Nitrogen. Guizhou Agricultural Sciences 11, 028.
559 Han, M., Gong, Y., Zhou, C., Zhang, J., Wang, Z. and Ning, K. (2016) Comparison and Interpretation of
560 Taxonomical Structure of Bacterial Communities in Two Types of Lakes on Yun-Gui plateau of China. Sci Rep 6.
561 Horppila, J. and Nurminen, L. (2003) Effects of submerged macrophytes on sediment resuspension and internal
562 phosphorus loading in Lake Hiidenvesi (southern Finland). Water research 37(18), 4468-4474.
563 Kang, C., Zhang, Y., Zhu, X., Liu, K., Wang, X., Chen, M., Wang, J., Chen, H., Hui, S. and Huang, L. (2016)
564 Healthy subjects differentially respond to dietary capsaicin correlating with specific gut enterotypes. The Journal of

565 Clinical Endocrinology and Metabolism. 101(12), 4681-4689.

566 Langfelder, P. and Horvath, S. (2012) Fast R functions for robust correlations and hierarchical clustering. Journal of
567 statistical software 46(11).

568 Langille, M.G., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J.A., Clemente, J.C., Burkepile,
569 D.E., Thurber, R.L.V. and Knight, R. (2013) Predictive functional profiling of microbial communities using 16S
570 rRNA marker gene sequences. Nature Biotechnology. 31(9), 814-821.

571 Lee, L.S., Carmosini, N., Sassman, S.A., Dion, H.M. and Sepulveda, M.S. (2007) Agricultural contributions of
572 antimicrobials and hormones on soil and water quality. Advances In Agronomy 93, 1-68.

573 Li, J., Zhang, J., Liu, L., Fan, Y., Li, L., Yang, Y., Lu, Z. and Zhang, X. (2015) Annual periodicity in planktonic
574 bacterial and archaeal community composition of eutrophic Lake Taihu. Scientific Reports 5.

575 Lindström, E.S., Kamst-Van Agterveld, M.P. and Zwart, G. (2005) Distribution of typical freshwater bacterial
576 groups is associated with pH, temperature, and lake water retention time. Applied and Environmental Microbiology
577 71(12), 8201-8206.

578 Lozupone, C. and Knight, R. (2005) UniFrac: a new phylogenetic method for comparing microbial communities.
579 Applied and Environmental Microbiology. 71(12), 8228-8235.

580 Ma, B., Wang, H., Dsouza, M., Lou, J., He, Y., Dai, Z., Brookes, P.C., Xu, J. and Gilbert, J.A. (2016) Geographic
581 patterns of co-occurrence network topological features for soil microbiota at continental scale in eastern China. The
582 ISME Journal.

583 Magurran, A.E. (2013) Measuring biological diversity, John Wiley & Sons.

584 Nielsen, P.H., Kragelund, C., Seviour, R.J. and Nielsen, J.L. (2009) Identity and ecophysiology of filamentous
585 bacteria in activated sludge. FEMS microbiology reviews 33(6), 969-998.

586 O'neil, J., Davis, T.W., Burford, M.A. and Gobler, C. (2012) The rise of harmful cyanobacteria blooms: the
587 potential roles of eutrophication and climate change. Harmful Algae 14, 313-334.

588 Pei, R., Kim, S.-C., Carlson, K.H. and Pruden, A. (2006) Effect of river landscape on the sediment concentrations
589 of antibiotics and corresponding antibiotic resistance genes (ARG). Water research 40(12), 2427-2435.

590 Pesaro, M. and Widmer, F. (2006) Identification and specific detection of a novel Pseudomonadaceae cluster
591 associated with soils from winter wheat plots of a long-term agricultural field experiment. Applied And
592 Environmental Microbiology 72(1), 37-43.

593 Pons, P. and Latapy, M. (2005) Computing communities in large networks using random walks. International
594 Symposium on Computer and Information Sciences, 284-293.

595 Porebski, S., Bailey, L.G. and Baum, B.R. (1997) Modification of a CTAB DNA extraction protocol for plants
596 containing high polysaccharide and polyphenol components. Plant molecular biology reporter 15(1), 8-15.

597 Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. and Glöckner, F.O. (2013) The
598 SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids
599 Research. 41(D1), D590-D596.

600 Rambaut, A. (2014) FigTree v1. 4.2. Edinburgh: University of Edinburgh.

601 Sand-jensen, K. (1998) Influence of submerged macrophytes on sediment composition and near-bed flow in
602 lowland streams. Freshwater biology 39(4), 663-679.

603 Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B.,
604 Parks, D.H. and Robinson, C.J. (2009) Introducing mothur: open-source, platform-independent,
605 community-supported software for describing and comparing microbial communities. Applied and Environmental
606 Microbiology. 75(23), 7537-7541.

607 Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S. and Huttenhower, C. (2011)
608 Metagenomic biomarker discovery and explanation. Genome Biology 12(6), R60.

609   Sengupta, S., Chattopadhyay, M.K. and Grossart, H.-P. (2013) The multifaceted roles of antibiotics and antibiotic
610   resistance in nature. Frontiers in Microbiology 4.
611   Shao, K., Gao, G., Wang, Y., Tang, X. and Qin, B. (2013) Vertical diversity of sediment bacterial communities in
612   two different trophic states of the eutrophic Lake Taihu, China. Journal of Environmental Sciences. 25(6),
613   1186-1194.
614   Sunagawa, S., Coelho, L.P., Chaffron, S., Kultima, J.R., Labadie, K., Salazar, G., Djahanschiri, B., Zeller, G.,
615   Mende, D.R. and Alberti, A. (2015) Structure and function of the global ocean microbiome. Science 348(6237),
616   1261359.
617   Sundaramanickam, A., Kumar, P.S., Kumaresan, S. and Balasubramanian, T. (2015) Isolation and molecular
618   characterization of multidrug-resistant halophilic bacteria from shrimp farm effluents of Parangipettai coastal
619   waters. Environmental Science and Pollution Research 22(15), 11700-11707.
620   Takamura N, K.Y., Fukushima M, Nakagawa M, KIM, B. H. (2003) Effects of aquatic macrophytes on water
621   quality and phytoplankton communities in shallow lakes. Ecological Research 18(4), 381-395.
622   Ter Braak, C.J. (1986) Canonical correspondence analysis: a new eigenvector technique for multivariate direct
623   gradient analysis. Ecology 67(5), 1167-1179.
624   Vázquez-Baeza, Y., Pirrung, M., Gonzalez, A. and Knight, R. (2013) EMPeror: a tool for visualizing
625   high-throughput microbial community data. Gigascience 2(1), 1.
626   Verdonschot, R., Keizer-vlek, H.E. and Verdonschot, P.F. (2011) Biodiversity value of agricultural drainage ditches:
627   a comparative analysis of the aquatic invertebrate fauna of ditches and small lakes. Aquatic Conservation: Marine
628   and Freshwater Ecosystems 21(7), 715-727.
629   Virtanen, R., Ilmonen, J., Paasivirta, L. and Muotka, T. (2009) Community concordance between bryophyte and
630   insect assemblages in boreal springs: a broad-scale study in isolated habitats. Freshwater Biology 54(8), 1651-1662.
631   Wang, H., Sangwan, N., Li, H.-Y., Su, J.-Q., Oyang, W.-Y., Zhang, Z.-J., Gilbert, J.A., Zhu, Y.-G., Ping, F. and
632   Zhang, H.-L. (2017a) The antibiotic resistome of swine manure is significantly altered by association with the
633   Musca domestica larvae gut microbiome. The ISME Journal 11(1), 100-111.
634   Wang, X. and Zhao, H. (2013) Isolation and Characterization of a Bacillus flexus Strain Used in Alkaline
635   Wastewater Treatment. Advanced Materials Research 750, 1381-1384.
636   Wang, Z., Du, Y., Yang, C., Liu, X., Zhang, J., Li, E., Zhang, Q. and Wang, X. (2017b) Occurrence and ecological
637   hazard assessment of selected antibiotics in the surface waters in and around Lake Honghu, China. Science of The
638   Total Environment 609(Supplement C), 1423-1432.
639   Williams, C.J., Frost, P.C., Morales-Williams, A.M., Larson, J.H., Richardson, W.B., Chiandet, A.S. and
640   Xenopoulos, M.A. (2016) Human activities cause distinct dissolved organic matter composition across freshwater
641   ecosystems. Global change biology 22(2), 613-626.
642   Williams, P., Whitfield, M., Biggs, J., Bray, S., Fox, G., Nicolet, P. and Sear, D. (2004) Comparative biodiversity of
643   rivers, streams, ditches and ponds in an agricultural landscape in Southern England. Biological conservation 115(2),
644   329-341.
645   Wood, S.A., Maier, M.Y., Puddick, J., Pochon, X., Zaiko, A., Dietrich, D.R. and Hamilton, D.P. (2017) Trophic
646   state and geographic gradients influence planktonic cyanobacterial diversity and distribution in New Zealand lakes.
647   FEMS Microbiology Ecology 93(2), fiw234.
648   Xu, H., Paerl, H.W., Qin, B., Zhu, G. and Gao, G. (2010) Nitrogen and phosphorus inputs control phytoplankton
649   growth in eutrophic Lake Taihu, China. Limnology and Oceanography 55(1), 420.
650   Yang, Y., Cao, X., Lin, H. and Wang, J. (2016) Antibiotics and Antibiotic Resistance Genes in Sediment of Honghu
651   Lake and East Dongting Lake, China. Microbial Ecology 72(4), 791-801.
652   Ye, W., Liu, X., Lin, S., Tan, J., Pan, J., Li, D. and Yang, H. (2009) The vertical distribution of bacterial and

653 archaeal communities in the water and sediment of Lake Taihu. FEMS Microbiology Ecology 70(2), 263-276.

654 Zhang, T., Ban, X., Wang, X., Cai, X., Li, E., Wang, Z., Yang, C., Zhang, Q. and Lu, X. (2017) Analysis of nutrient
655 transport and ecological response in Honghu Lake, China by using a mathematical model. Science of The Total
656 Environment 575, 418-428.

657 Zhang, X. (1998) On the estimation of biomass of submerged vegetation using Landsat thematic mapper (TM)
658 imagery: a case study of the Honghu Lake, PR China. International Journal Of Remote Sensing 19(1), 11-20.

659

660 Declarations

661 Authors' contributions

662 The whole study was designed by ZW and KN. MZH, JQZ and ZW collected samples. MZH,

663 CYC, QY and HZ performed DNA extraction and sequencing. MZH, CYZ, MD and HJL analyzed

664 the data. MZH, CYZ, MD, HJL, JG, ZW, and KN wrote the initial draft of the manuscript. All revised

665 the manuscript.

666 Funding

672 Availability of data and material

673 All sequencing data for the 14 water samples and the 14 sediment samples were deposited into

674 NCBI's Sequence Read Archive (SRA) database under the Bioproject number PRJNA352457.

675 Competing financial interests

676 The authors declare no competing financial interests.

677 Ethics approval and consent to participate

678   Not applicable.

Consent for publication

680   Not applicable.

Acknowledgements

682   Not applicable.

683

684 **Figure Legends**

685 **Figure 1. Geographic distribution of all sampling sites in Honghu lake.** L: Lake; P: Pond; R:

686 River. **(A)** Definitions of the various sampling strategies. **(B)** Locations of sampling sites and the

687 distribution of the sampling medium collected at each site.

688

689 **Figure 2. Taxonomic composition and relative abundances of microbial taxa in water and**

690 **sediment samples. (A)** Taxonomic composition of each sample at the phylum level. 'Other'

691 represents all phyla not in the top 13 phyla. **(B)** Bar plot highlighting differences between water and

692 sediment samples at phylum level. ** represents p-values <0.01 and determined by a student t-test.

693 **(C)** Bar plot highlighting differences in water samples at the phylum level between impacted and

694 less-impacted groups. * represents p-values <0.05 and determined by a student t-test. **(D)** Bar plot

695 highlighting differences in sediment samples at the phylum level between impacted and

696 less-impacted groups. * represents p-values <0.05 and determined by a student t-test. ** represents

697 p-values <0.01 and determined by a student t-test.

698

699 **Figure 3. Core- and Pan-OTUs of water and sediment samples from Honghu lake.** Flower plots

700 showing the number of sample-specific OTUs (in the petals) and Core-OTUs (in the center) for **(A)**

701 all samples, **(B)** all water samples, and **(C)** all sediment samples. OTU accumulation curves for

702 Pan-OTUs (upper) and Core-OTUs (lower) for **(D)** all samples, **(E)** all water samples, and **(F)** all

703 sediment samples from Honghu lake.

704

705 **Figure 4. PCoA plots and UPGMA-based clustering of water and sediment microbial**

706 **communities.** Unweighted UniFrac dissimilarity matrix scores for all samples visualized in a PCoA

707 plot to demonstrate the dissimilarity of the microbial community structure between samples by **(A)**

708   sampling medium (water vs. sediment) and by **(B)** impacted and less-impacted groups. **(C)**

709   UPGMA-based clustering tree of microbial communities using an unweighted UniFrac distance

710   matrix. The green and pink font represent less-impacted and impacted groups, respectively. The blue

711   and red bars mark water and sediment samples, respectively.

712

713   **Figure 5. Canonical correspondence analysis plots of physicochemical properties and antibiotic**

714   **data driving water and sediment microbial community structure.** Physicochemical properties of

715   **(A)** water samples **(B)** and sediment samples from Honghu lake, and antibiotic data for **(C)** water

716   samples and **(D)** sediment samples from Honghu lake. *** represents *p*-values <0.001, ** represents

717   *p*-values <0.01 and * represents *p*-values <0.05.

718

719   **Figure 6. Environmental drivers of microbial community composition in water and sediment**

720   **samples.** Pairwise comparisons of **(A)** water and **(B)** sediment physicochemical properties with

721   taxonomic and functional composition data. Color gradient represents Pearson's Correlations

722   Coefficients. Pairwise comparisons of **(C)** water and **(D)** sediment antibiotic concentration data with

723   taxonomic and functional composition data. In all figures, the varying circle size represents the

724   absolute value of the Pearson Correlation Coefficient between the two factors, the bar along the

725   y-axis represents the value of the Pearson Correlation Coefficients, and the edge width represents the

726   Mantel's *r* statistic value for distance correlations and the edge color denotes the statistical

727   significance based on 9,999 permutations. Abbreviations of physicochemical properties and

728   antibiotics are listed in sub-section '***Physicochemical characterization and antibiotic analysis***'.

729

730   **Figure 7. Biomarkers analysis of water and sediment microbial communities from impacted**

731   **and less-impacted (control) sites. (A)** Differentially abundant taxa of water samples; **(B)**

732   Cladogram showing the phylogenetic structure of the microbiota from water samples; **(C)**

733   Differentially abundant taxa of sediment samples; **(D)** Cladogram showing the phylogenetic structure

734   of the microbiota from sediment samples.

735

736   **Figure 8. Co-occurrence network interactions of Honghu lake microbes in water and sediment**

737   **samples.** Network nodes represent OTUs and edges are represented as interactions between OTUs.

738   All edges represent positive, strong (Spearman's $\rho > 0.8$), and significant (p-value<0.001) correlations.

739   The size of each node is proportional to the node degree. Networks of **(A)** water and **(B)** sediment

740   samples displaying co-occurrence patterns of OTUs grouped at the phylum level. Modules were

741   identified using the WalkTrap community detection algorithm in **(C)** water and **(D)** sediment samples.

742   Water **(E)** and **(F)** sediment networks investigating the effect of long-term agricultural activities

743   wherein each node was colored as function of its relative abundance in impacted and less-impacted

744   sites.

745

(A)
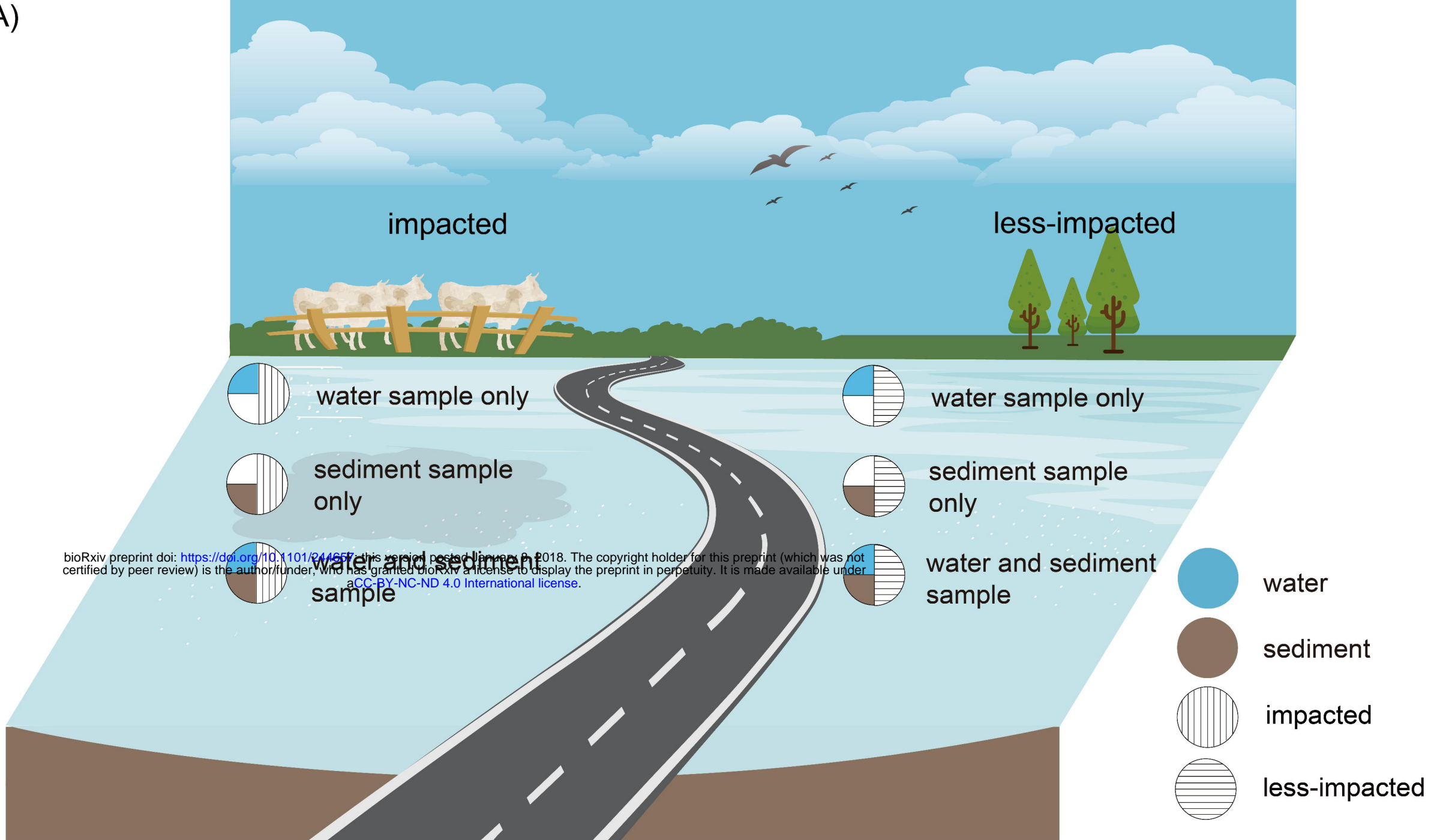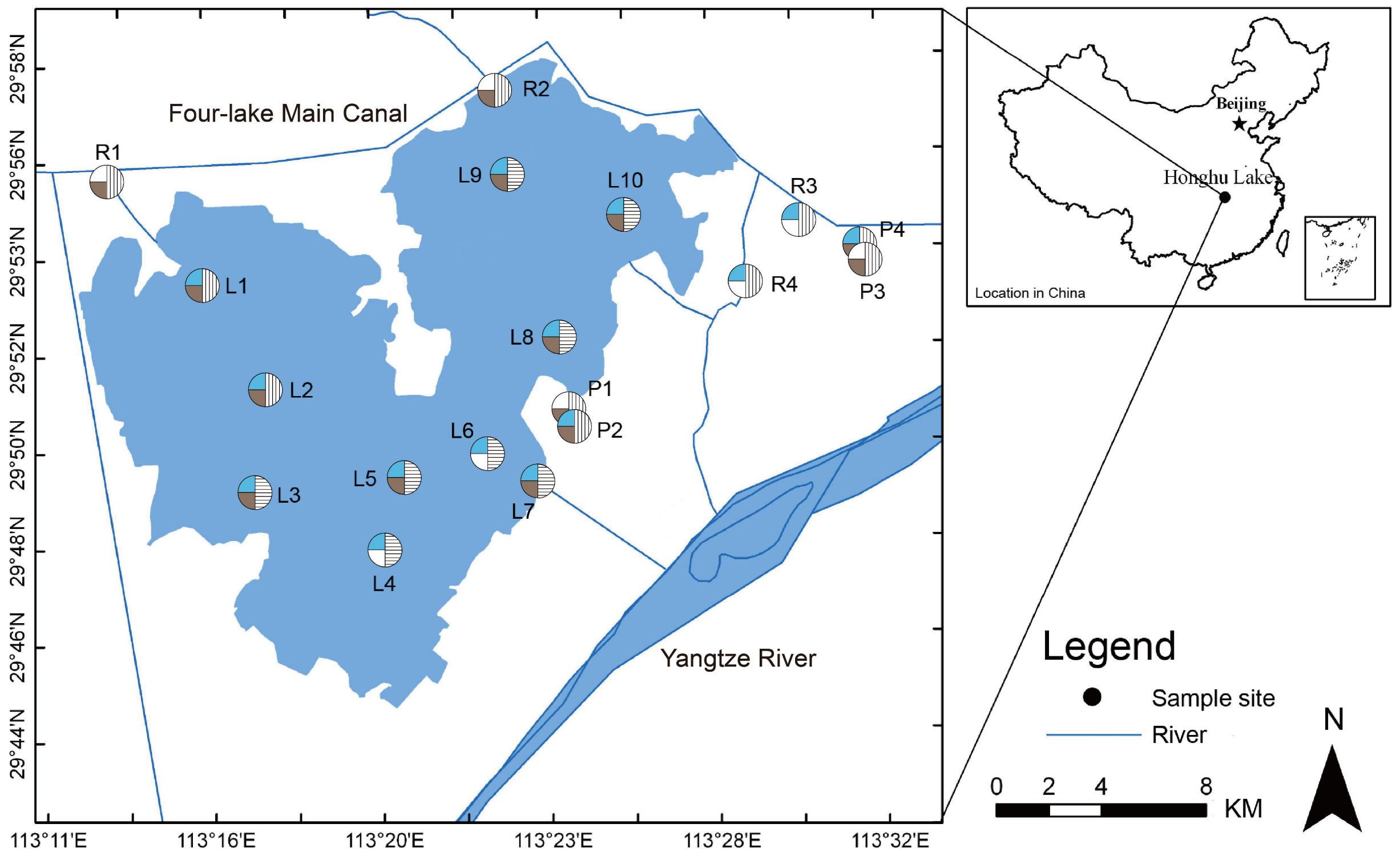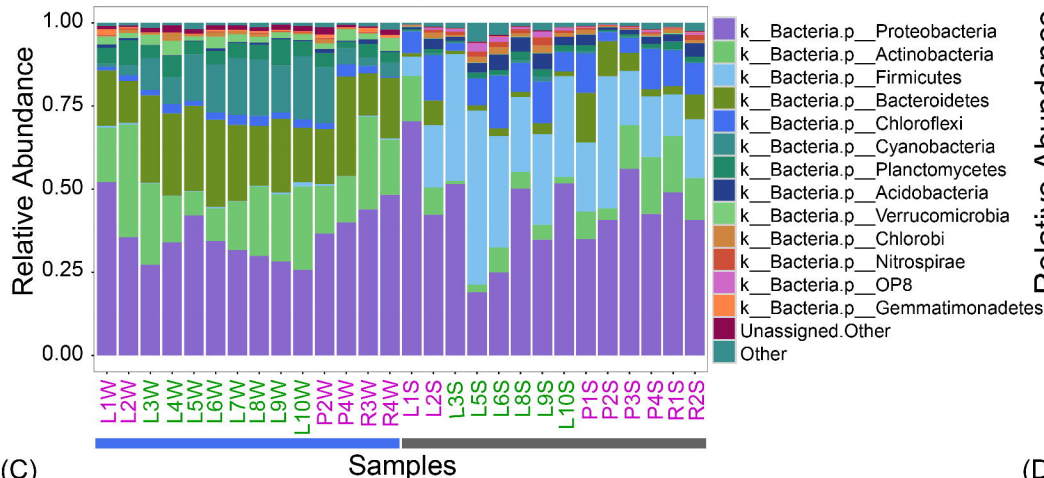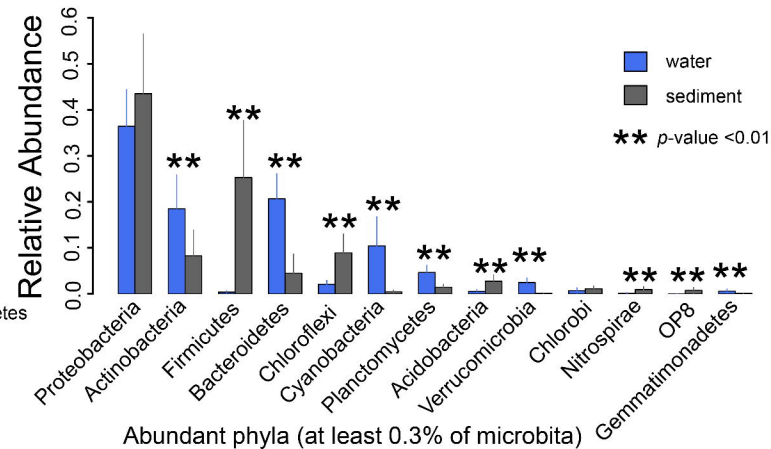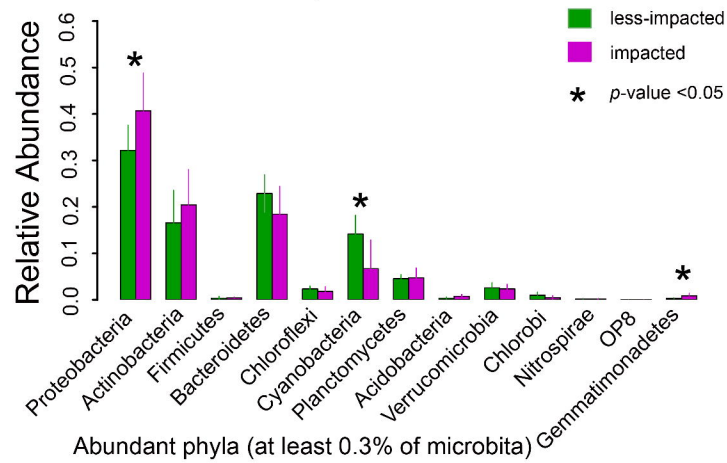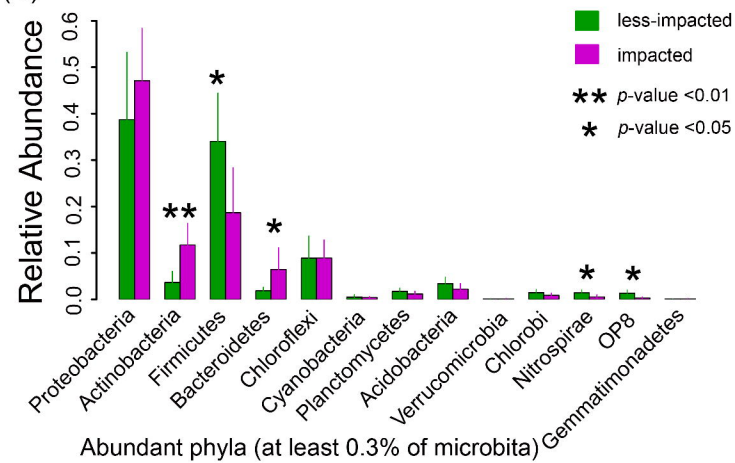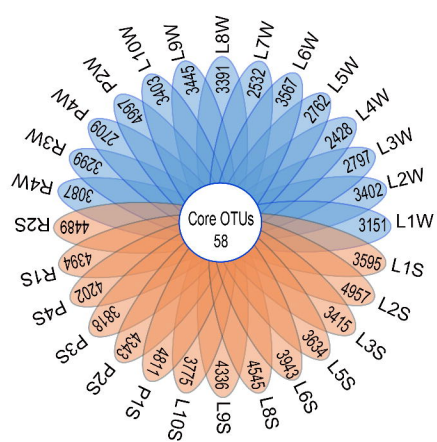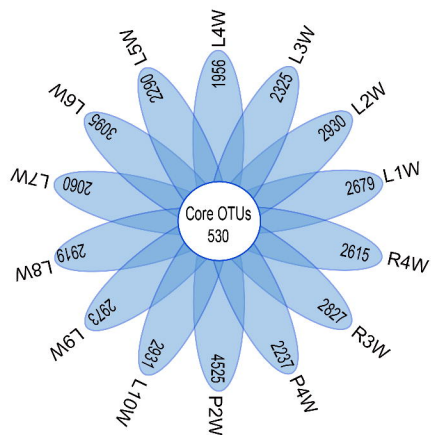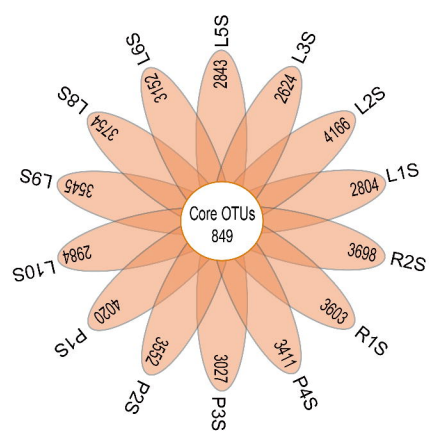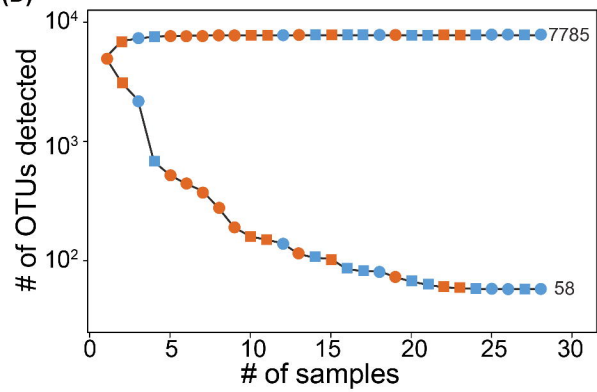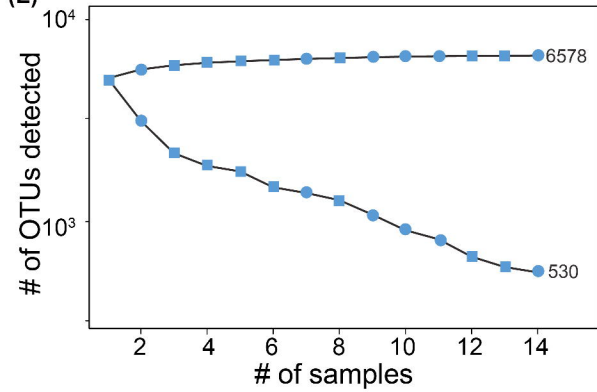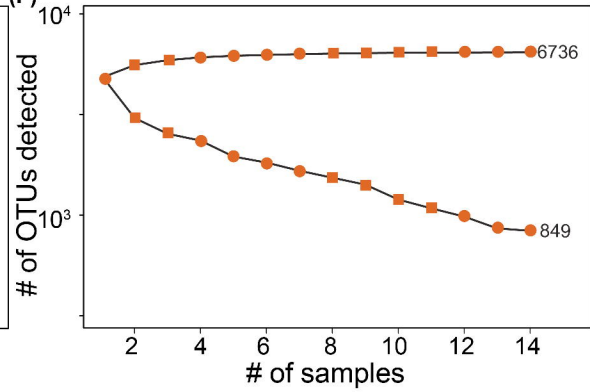
impacted   less-impacted

water sample only   water sample only

sediment sample
only   sediment sample
only

water and sediment
sample

water   sediment   sample

water

sediment

impacted

less-impacted

(B)

Four-lake Main Canal

R1

R2

L9   L10

R3

P4

L1   R4

P3

L8

L2   P1

L6   P2

L5

L3   L7

L4

Yangtze River

Four-lake Main Canal

Beijing

Honghu Lake

Location in China

Legend

● Sample site

— River

0 2 4   8

KM

N

29°58'N
29°56'N
29°53'N
29°52'N
29°50'N
29°48'N
29°46'N
29°44'N

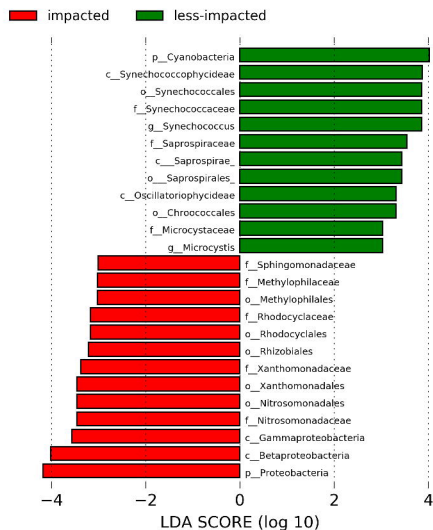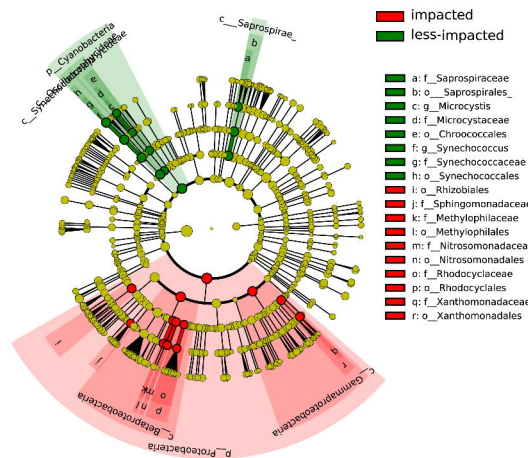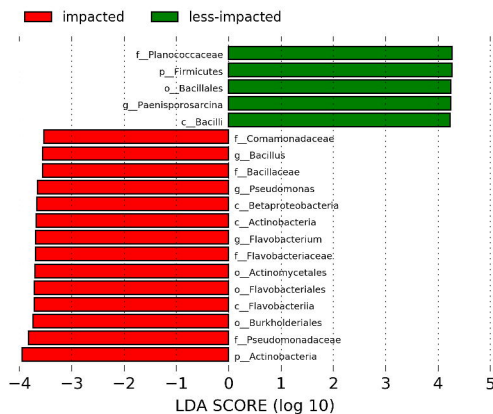113°11'E   113°16'E   113°20'E   113°23'E   113°28'E   113°32'E

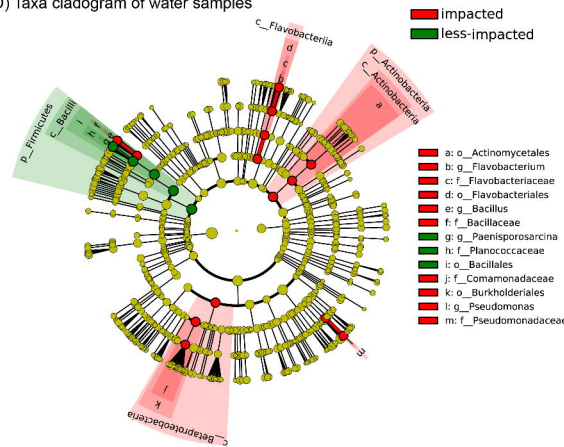(A) Differently abundant taxa of water samples
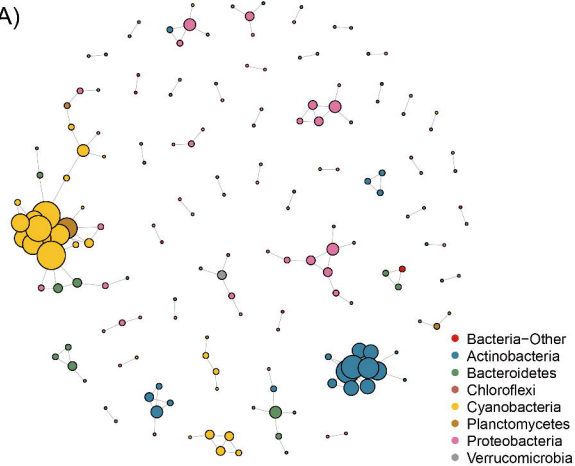
(B) Taxa cladogram of water samples
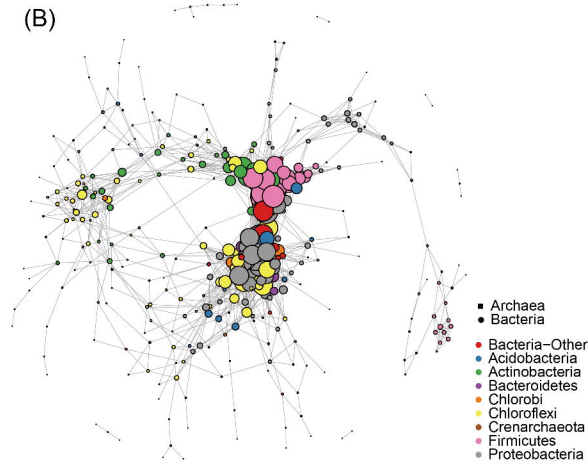
(C) Differently abundant taxa of sediment samples

(D) Taxa cladogram of water samples

(A)

Bacteria−Other
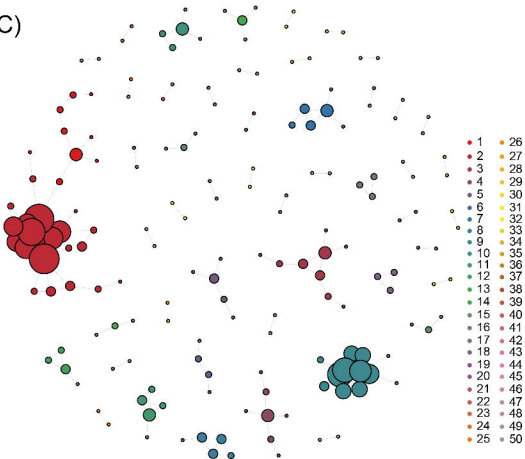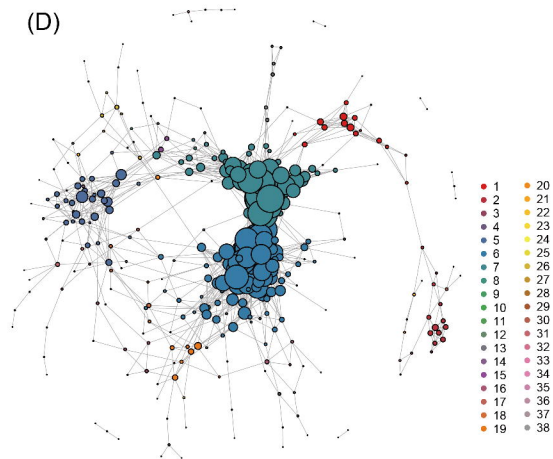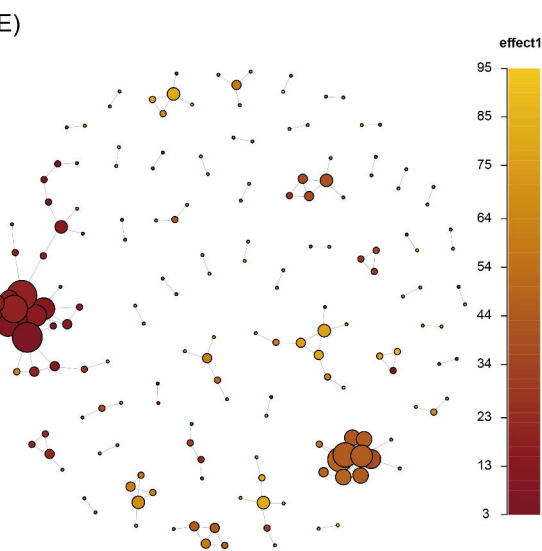Actinobacteria
Bacteroidetes
Chloroflexi
Cyanobacteria
Planctomycetes
Proteobacteria
Verrucomicrobia

(B)

■ Archaea
● Bacteria

Bacteria−Other
Acidobacteria
Actinobacteria
Bacteroidetes
Chlorobi
Chloroflexi
Crenarchaeota
Firmicutes
Proteobacteria

(C)

1	26
2	27
3	28
4	29
5	30
6	31
7	32
8	33
9	34
10	35
11	36
12	37
13	38
14	39
15	40
16	41
17	42
18	43
19	44
20	45
21	46
22	47
23	48
24	49
25	50

(D)

1	20
2	21
3	22
4	23
5	24
6	25
7	26
8	27
9	28
10	29
11	30
12	31
13	32
14	33
15	34
16	35
17	36
18	37
19	38

(E)

effect1

95
85
75
64
54
44
34
23
13
3

(F)

■ Archaea
● Bacteria

effect1

98
87
77
66
55
45
34
23
13
2