

1 **Bacterial Community Structure Correlates with *Legionella***  
2 ***pneumophila* Colonization of New York City High Rise Building**  
3 **Premises Plumbing Systems**

4

5 **Supplementary Information**

6

7

8

9

10

11 Xiao Ma <sup>a</sup>, David Pierre <sup>a</sup>, Kyle Bibby <sup>b,c,1</sup>, Janet E. Stout <sup>a,b,\*</sup>

12

13

14

15

16

17

18

19

20 <sup>a</sup> Special Pathogens Laboratory, Pittsburgh, PA 15219, USA

21 <sup>b</sup> Department of Civil and Environmental Engineering, University of Pittsburgh, Pittsburgh, PA

22 15261, USA

23 <sup>c</sup> Department of Computational and Systems Biology, University of Pittsburgh Medical School,

24 Pittsburgh, PA 15261, USA

25

26

27

28 \*Corresponding Author: Janet E. Stout, 1401 Forbes Ave, Suite 401, Pittsburgh, PA 15219,

29 [jstout@specialpathogenslab.com](mailto:jstout@specialpathogenslab.com), 412-281-5335

30

31 <sup>1</sup> (Current affiliation) Department of Civil and Environmental Engineering and Earth Sciences,

32 University of Notre Dame, Notre Dame, IN 46556, USA

33

34

35

36

37

38

39

40

41 SI-Table 1. RDP taxonomy classification results of pure culture *Legionella* DNA

Pure culture	Assigned taxa	Percentage assigned
<i>Legionella pneumophila</i>	<i>Legionella pneumophila</i>	99.80%
	<i>Legionella</i> (unclassified species)	0.02%
	<i>Legionella cherrii</i>	0.02%
	Not assigned as Legionellaceae	0.16%
<i>Legionella micdadei</i>	<i>Legionella micdadei</i>	99.90%
	<i>Legionella cherrii</i>	0.02%
	Not assigned as Legionellaceae	0.08%
<i>Legionella anisa</i>	<i>Legionella</i> (unclassified species)	9.23%
	<i>Legionella cherrii</i>	90.40%
	<i>Legionella gratiana</i>	0.03%
	<i>Legionella steelei</i>	0.08%
	<i>Legionella steigerwaltii</i>	0.06%
	Not assigned as Legionellaceae	0.21%

42

43

44 SI-Table 2. Number of samples positive or negative for *Legionella pneumophila*

		<i>L. pneumophila</i> culture	
		Positive	Negative
<i>L. pneumophila</i> 16S	Positive	16	7
	Negative	4	54

45

46

47

48

49

50

51

52

53

54

55

SI-Table 3. Least discriminant effect size (LEfSe) analysis to class level of bacterial community in cold water (CW) and hot water (HW)

Bacterial taxa	(log 10) Highest Mean Value	Temperature group	LDA Score (log 10)	P-Value
k__Archaea.p__Euryarchaeota	1.36	CW	3.98	0.019786
k__Archaea.p__Euryarchaeota.c__Methanobacteria	1.36	CW	4.07	0.019786
k__Bacteria.p__Thermi__	4.09	HW	3.78	0.009107
k__Bacteria.p__Thermi.c__Deinococci	4.09	HW	3.78	0.009107
k__Bacteria.p__Acidobacteria	4.92	HW	4.48	1.97E-05
k__Bacteria.p__Acidobacteria.c__Chloracidobacteria__	4.86	HW	4.48	8.13E-05
k__Bacteria.p__Acidobacteria.c__Acidobacteriia	1.49	CW	3.10	0.004079
k__Bacteria.p__Acidobacteria.c__Holophagae	4.19	CW	3.62	0.007257
k__Bacteria.p__Acidobacteria.c__Solibacteres	4.00	HW	3.63	2.24E-05
k__Bacteria.p__Armatimonadetes	0.92	CW	3.46	0.019786
k__Bacteria.p__Bacteroidetes	5.14	HW	4.73	1.05E-05
k__Bacteria.p__Bacteroidetes.c__	4.68	HW	4.36	0.047202
k__Bacteria.p__Bacteroidetes.c__Saprospirae__	4.90	HW	4.48	2.49E-05
k__Bacteria.p__Chlorobi.c__OPB56	3.29	HW	2.95	0.028952
k__Bacteria.p__Chloroflexi	3.01	HW	2.57	0.009265
k__Bacteria.p__Chloroflexi.c__Ellin6529	2.98	HW	2.67	0.008015
k__Bacteria.p__Firmicutes.c__Bacilli	3.73	CW	3.27	0.012176
k__Bacteria.p__Gemmatimonadetes	4.47	HW	4.10	6.17E-05
k__Bacteria.p__Gemmatimonadetes.c__Gemm_1	4.47	HW	4.11	2.42E-05
k__Bacteria.p__Planctomycetes	4.35	HW	3.86	0.047075
k__Bacteria.p__Planctomycetes.c__Phycisphaerae	1.40	CW	3.17	0.019786
k__Bacteria.p__Planctomycetes.c__Planctomycetia	4.35	HW	3.86	0.047075
k__Bacteria.p__Proteobacteria	5.96	CW	5.02	3.14E-06
k__Bacteria.p__Proteobacteria.c__Alphaproteobacteria	5.67	CW	4.99	0.010824
k__Bacteria.p__Proteobacteria.c__Betaproteobacteria	5.59	HW	4.78	0.04708
k__Bacteria.p__Proteobacteria.c__Gammaproteobacteria	5.22	CW	4.87	1.88E-05
k__Bacteria.p__SBR1093	1.46	HW	2.28	0.030967
k__Bacteria.p__SBR1093.c__	1.44	HW	2.29	0.040964

57

58 SI-Table 4. ADONIS analysis of correlation between bacterial community variation and water quality parameters

Parameters	Unweighted UniFrac				Weighted UniFrac			
	Hot water		Cold water		Hot water		Cold water	
	R <sup>2</sup>	p	R <sup>2</sup>	p	R <sup>2</sup>	p	R <sup>2</sup>	p
Temperature	0.029	0.01	0.061	0.038	0.03	0.174	0.092	0.04
pH	0.028	0.017	0.067	0.01	0.015	0.54	0.045	0.42
Free chlorine	0.035	0.001	0.049	0.22	0.051	0.012	0.13	0.007
Fe	0.031	0.015	0.04	0.61	0.044	0.026	0.02	0.87
Ca	0.015	0.75	0.067	0.024	0.021	0.253	0.077	0.079
Mg	0.017	0.54	0.072	0.011	0.0088	0.828	0.048	0.367
Zn	0.02	0.31	0.054	0.12	0.011	0.721	0.03	0.65
Pb	0.02	0.27	0.039	0.65	0.016	0.449	0.017	0.932
Mn	0.035	0.003	0.037	0.75	0.042	0.031	0.035	0.58
Cu	0.032	0.01	0.047	0.29	0.047	0.016	0.038	0.52
TOC	0.12	0.003	NA	NA	0.069	0.461	NA	NA

59 NA: not applicable. TOC was only measured in hot water return line samples; ADONIS analysis of correlation between TOC and  
60 bacterial community variation were only conducted in hot water return line samples.

61

62

63

64

65

66

67

68

SI-Table 5. Least discriminant effect size (LEfSe) analysis to class level of cold water bacterial community in *L. pneumophila* positive and negative samples

Bacterial taxa	(log 10) Highest Mean Value	<i>L. pneumophila</i> positivity	LDA Score (log 10)	P-Value
k_Bacteria.p_Thermi_	4.04	P	4.15	0.020409
k_Bacteria.p_Thermi_.c_Deinococci	4.04	P	4.12	0.020409
k_Bacteria.p_Acidobacteria.c_Chloracidobacteria_	4.81	P	4.47	0.029547
k_Bacteria.p_Acidobacteria.c_Solibacteres	3.91	P	3.82	0.020279
k_Bacteria.p_Bacteroidetes	5.14	P	4.81	0.03001
k_Bacteria.p_Bacteroidetes.c_Saprospirae_	4.99	P	4.67	0.022324
k_Bacteria.p_Cyanobacteria	4.24	P	3.99	0.039807
k_Bacteria.p_Cyanobacteria.c_4C0d_2	4.24	P	3.99	0.039807
k_Bacteria.p_Gemmatimonadetes	4.53	P	4.28	0.014655
k_Bacteria.p_Gemmatimonadetes.c_Gemm_1	4.53	P	4.27	0.008055
k_Bacteria.p_Proteobacteria	5.97	N	5.08	0.039807
k_Bacteria.p_Proteobacteria.c_Deltaproteobacteria	4.56	P	4.22	0.022211

69

SI-Table 6. Least discriminant effect size (LEfSe) analysis to class level of hot water bacterial community in *L. pneumophila* positive and negative samples

Bacterial taxa	(log 10) Highest Mean Value	<i>L. pneumophila</i> positivity	LDA Score (log 10)	P-Value
k_Bacteria.p_Acidobacteria.c_Chloracidobacteria_	4.98	P	4.25	0.031031
k_Bacteria.p_Acidobacteria.c_Solibacteres	4.11	N	3.71	0.032804
k_Bacteria.p_Actinobacteria.c_Actinobacteria	3.97	P	3.49	0.035093
k_Bacteria.p_Bacteroidetes	5.39	P	4.85	0.035104
k_Bacteria.p_Bacteroidetes.c_Saprospirae_	5.06	P	4.45	0.021183
k_Bacteria.p_Bacteroidetes.c_Bacteroidia	1.32	N	3.80	0.033066
k_Bacteria.p_Gemmatimonadetes	4.52	P	4.19	0.010471
k_Bacteria.p_Gemmatimonadetes.c_Gemm_1	4.52	P	4.19	0.010151
k_Bacteria.p_Proteobacteria	5.87	N	4.90	0.017762
k_Bacteria.p_SBR1093.c	1.60	N	3.98	0.023408

70

71

72 SI-Table 7. Spearman's rank-order correlation coefficient between relative abundance of bacteria classes associated with *L.*

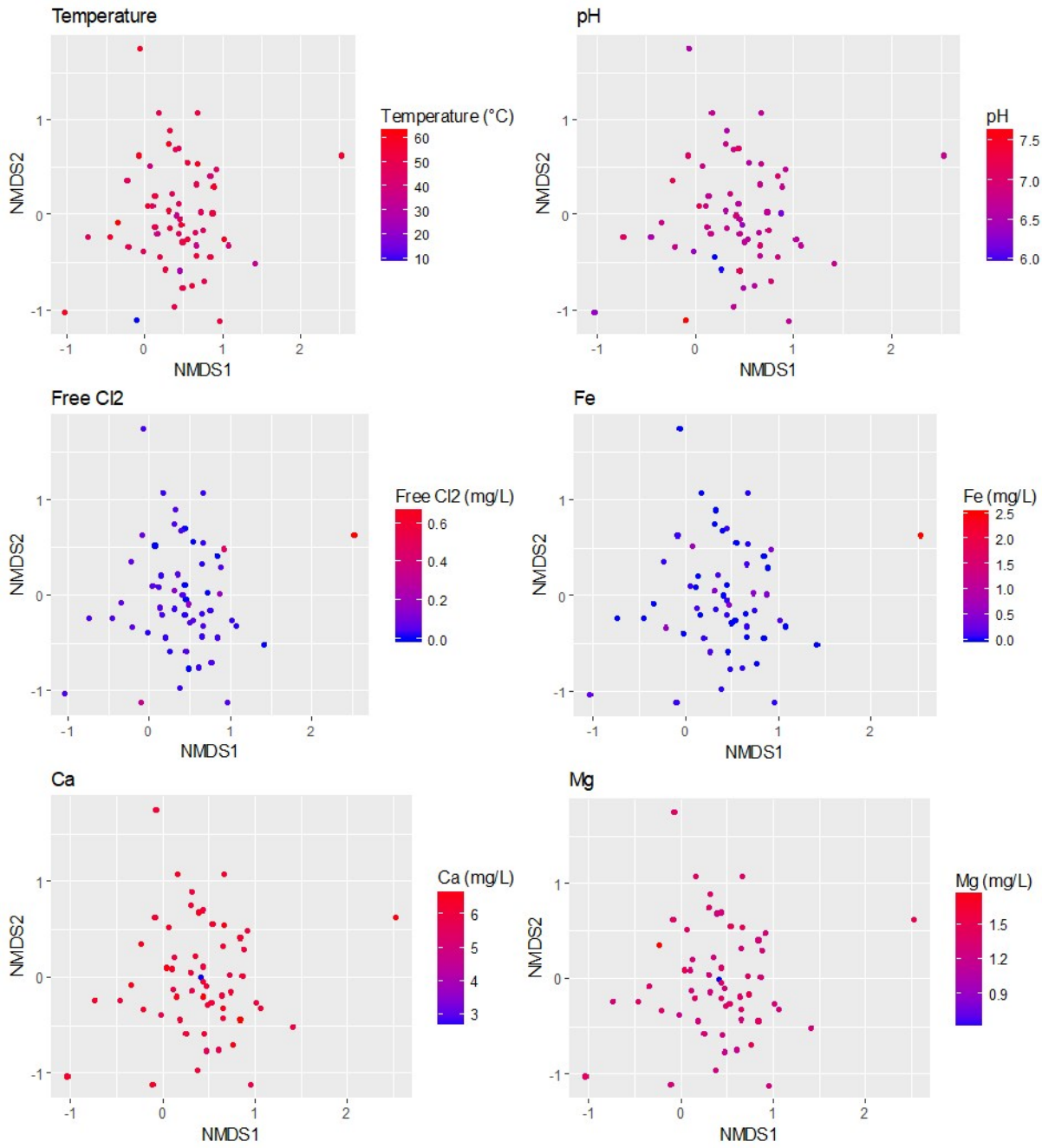
	Temperature	pH	Free Cl <sub>2</sub>	Fe	Ca	Mg	Zn	Pb	Mn	Cu	TOC
<b>(Hot water)</b>											
Bacteroidia	0.357 (0.006)					0.361 (0.005)				-0.292 (0.025)	
Solibacteres									-0.439 (0.001)		-0.612 (0.015)
Actinobacteria	-0.309 (0.017)										
Gemm-1							-0.264 (0.043)				
Chloracidobacteria										0.287 (0.028)	
Saprospirae	-0.333 (0.01)			-0.274 (0.036)							
<i>L. pneumophila</i>		0.32 (0.014)					-0.361 (0.014)			-0.347 (0.007)	
<b>(Cold water)</b>											
4C0d-2					0.524 (0.012)						
<i>L. pneumophila</i>				-0.426 (0.048)							

73 *pneumophila* and water quality parameters; p-values are shown in ( )

74 Note: Coefficients with p-value > 0.05 are not shown

75

76



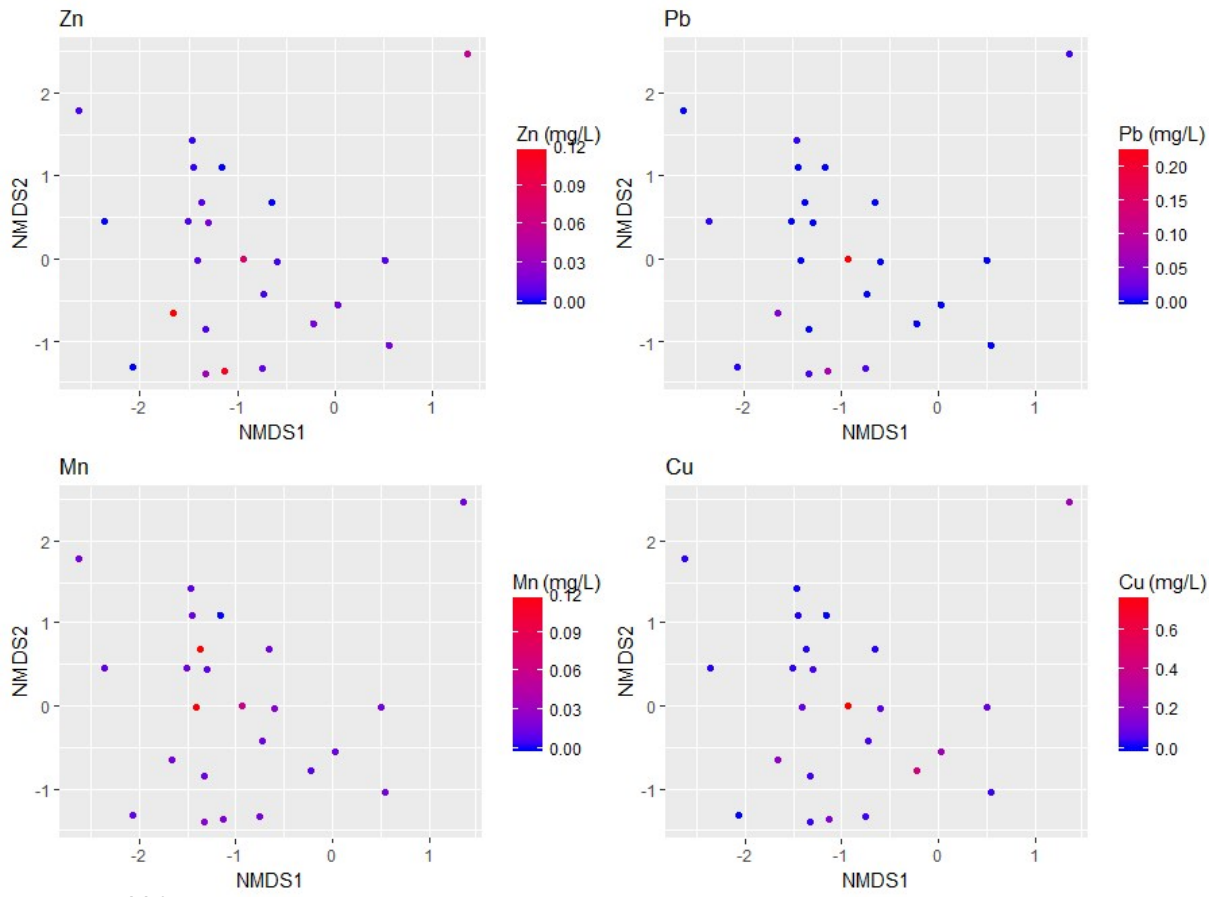
95

96

97

98 SI-Figure 1. Non-metric Multidimensional Scaling (NMDS) plot based on unweighted UniFrac  
99 distances showing hot water bacterial community in relation with water quality parameters

100



113

114

115

116

117 SI-Figure 1 (Continued). Non-metric Multidimensional Scaling (NMDS) plot based on  
118 unweighted UniFrac distances showing hot water bacterial community in relation with water  
119 quality parameters

120

121

122

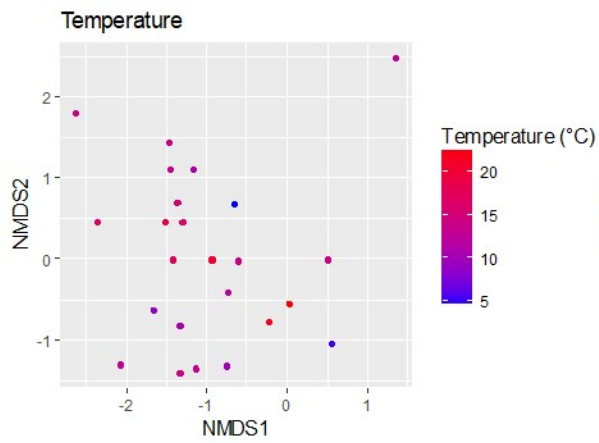
123

124

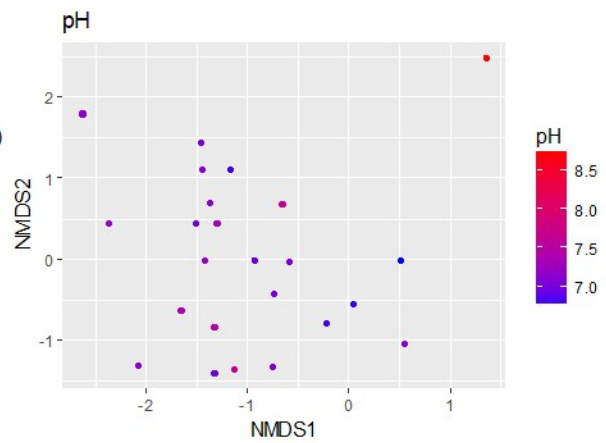


125

126



127



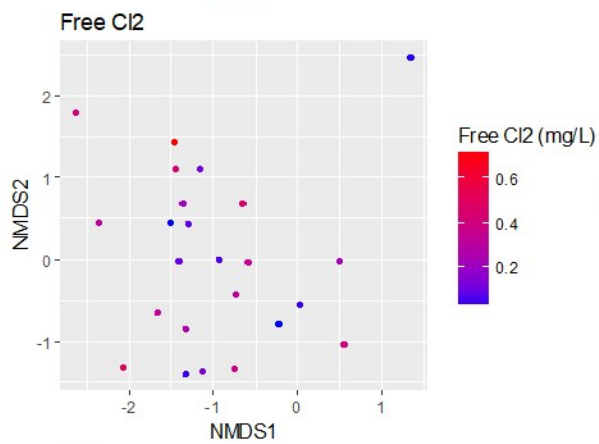
128

129

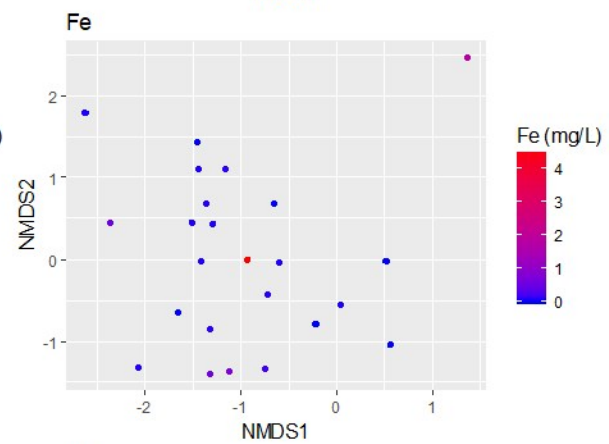
130

131

132



133



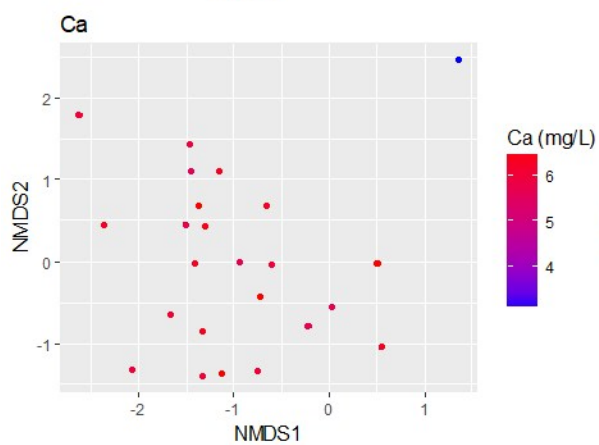
134

135

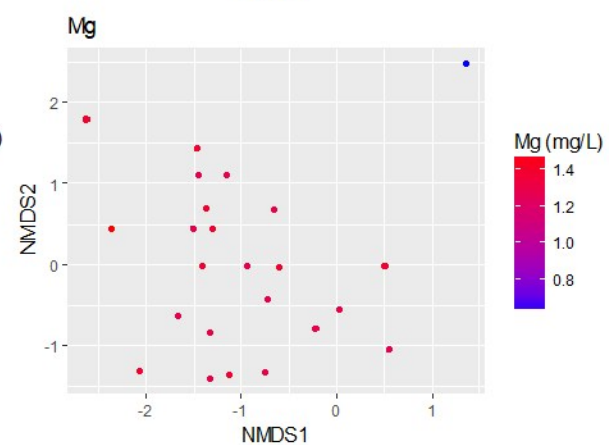
136

137

138



139



140

141

142

143

144

145

146

147 SI-Figure 2. Non-metric Multidimensional Scaling (NMDS) plot based on unweighted UniFrac  
148 distances showing cold water bacterial community in relation with water quality parameters

149

150

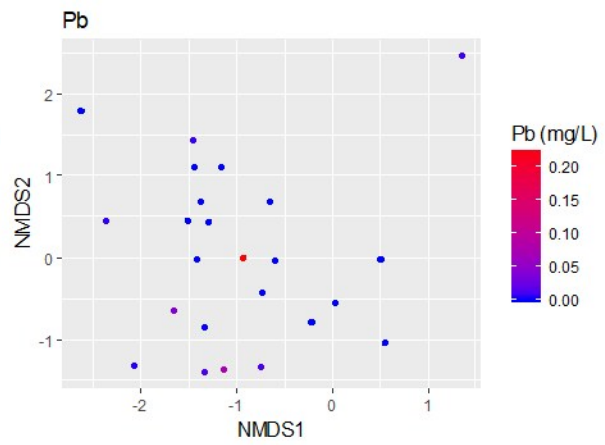
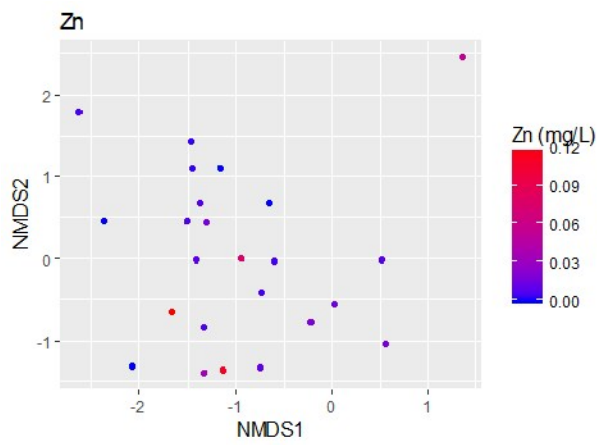
151

152

153

154

155



156

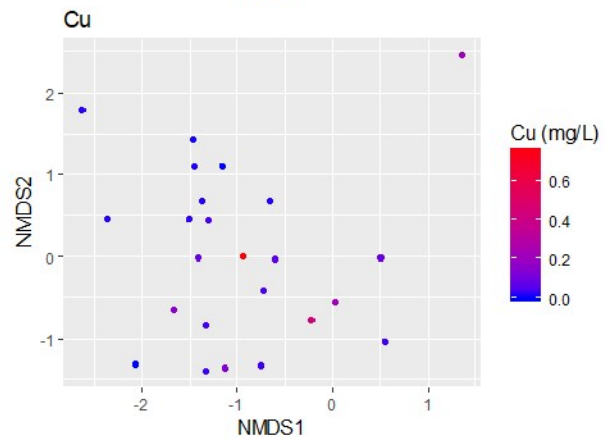
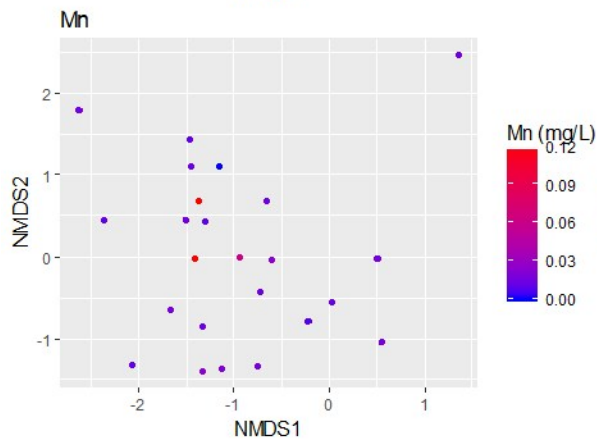
157

158

159

160

161



162

163

164

165 SI-Figure 2 (Continued). Non-metric Multidimensional Scaling (NMDS) plot based on  
166 unweighted UniFrac distances showing cold water bacterial community in relation with water  
167 quality parameters

168

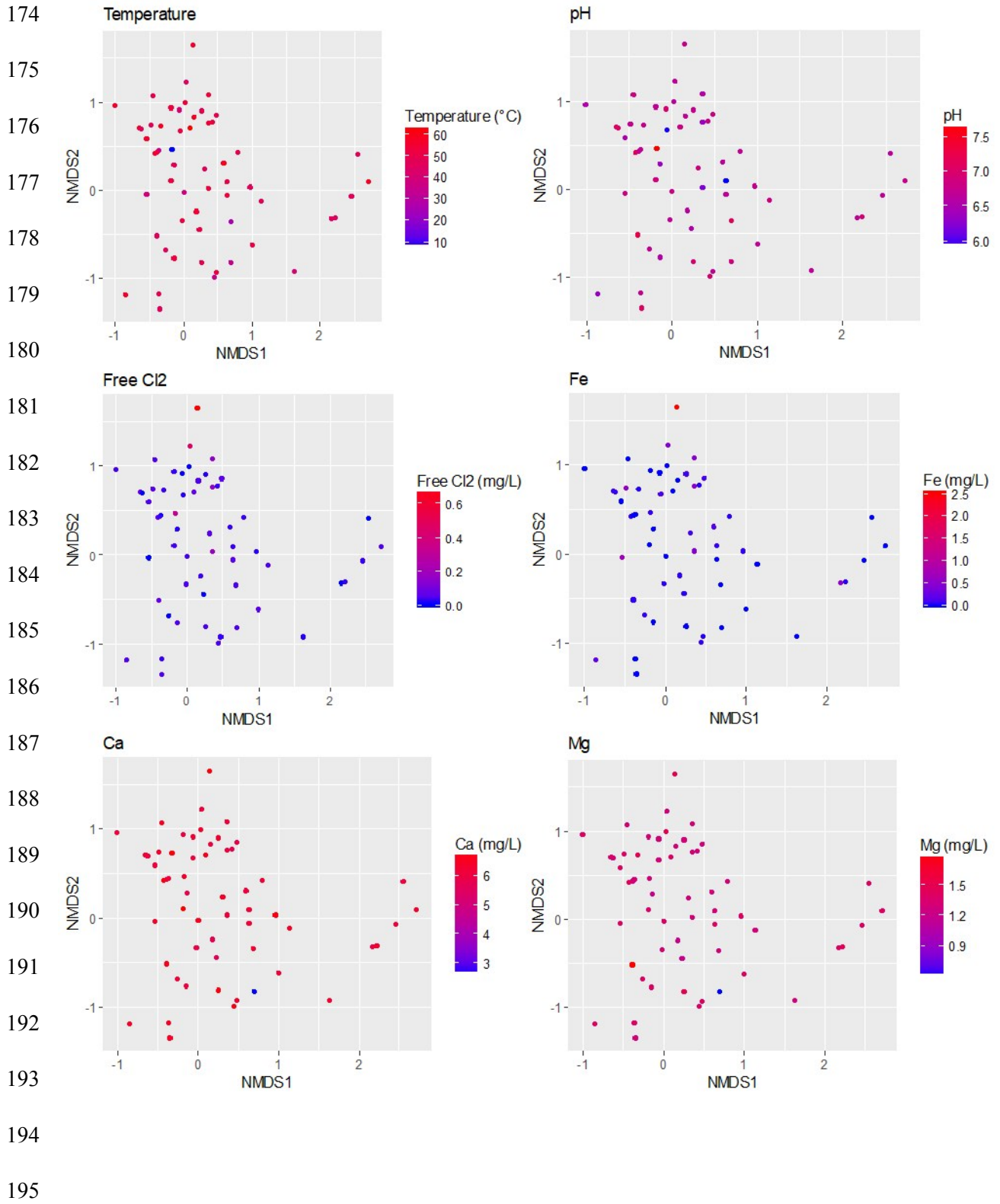
169

170

171

172

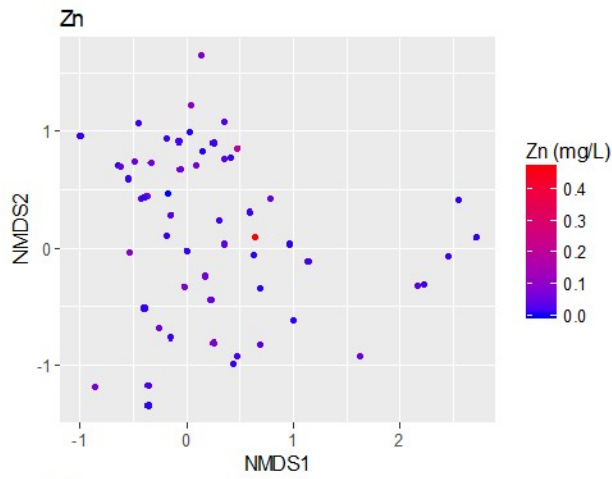
173



196 SI-Figure 3. Non-metric Multidimensional Scaling (NMDS) plot based on weighted UniFrac  
 197 distances showing hot water bacterial community in relation with water quality parameters

198  
199

200



201

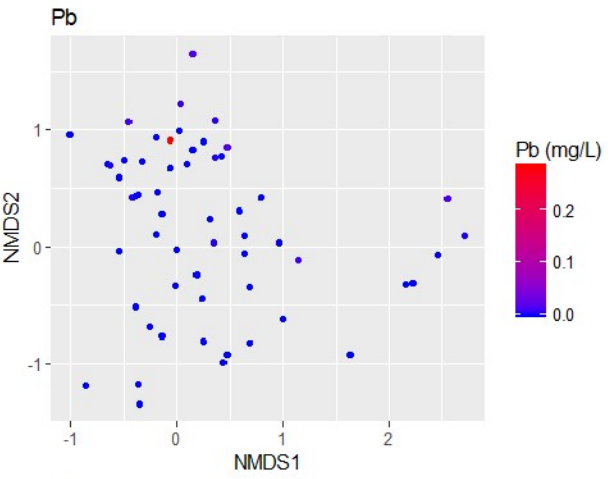
202

203

204

205

206



207

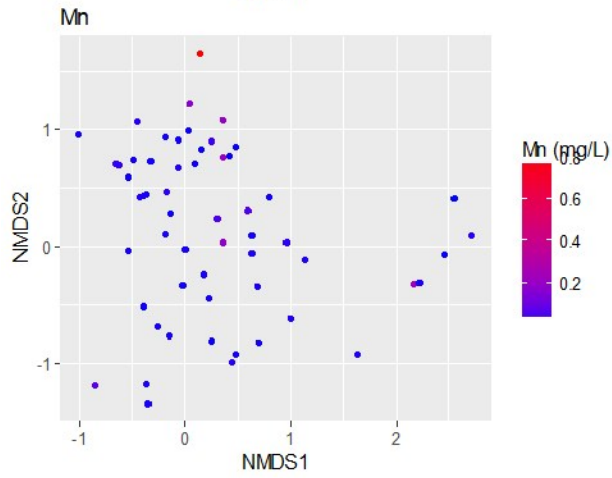
208

209

210

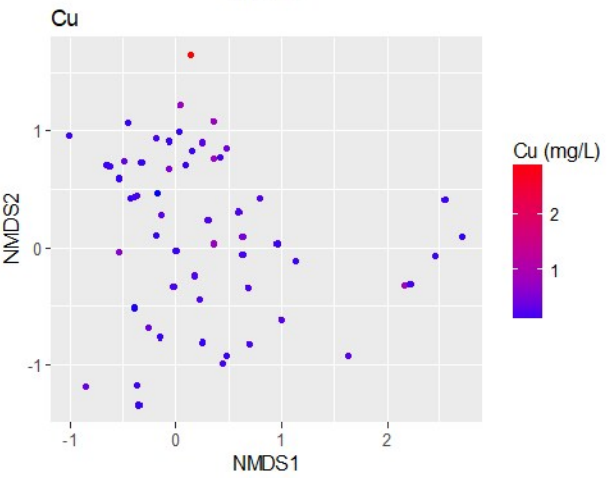
211

212



213

214



215 SI-Figure 3 (Continued). Non-metric Multidimensional Scaling (NMDS) plot based on weighted  
216 UniFrac distances showing hot water bacterial community in relation with water quality  
217 parameters

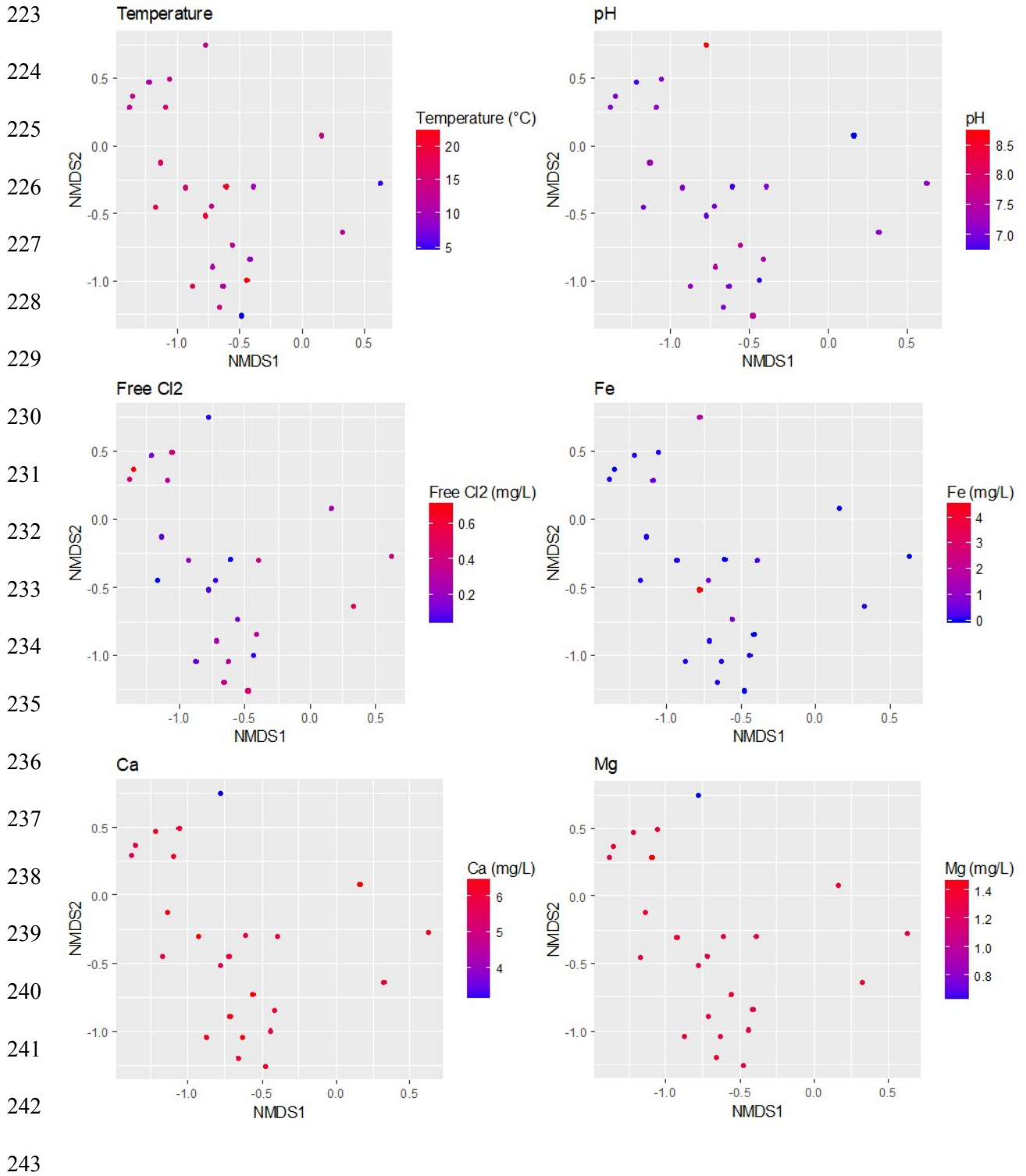
218

219

220

221

222



244 SI-Figure 4. Non-metric Multidimensional Scaling (NMDS) plot based on weighted UniFrac  
 245 distances showing cold water bacterial community in relation with water quality parameters  
 246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263 SI-Figure 4 (Continued). Non-metric Multidimensional Scaling (NMDS) plot based on weighted  
264 UniFrac distances showing cold water bacterial community in relation with water quality  
265 parameters  
266

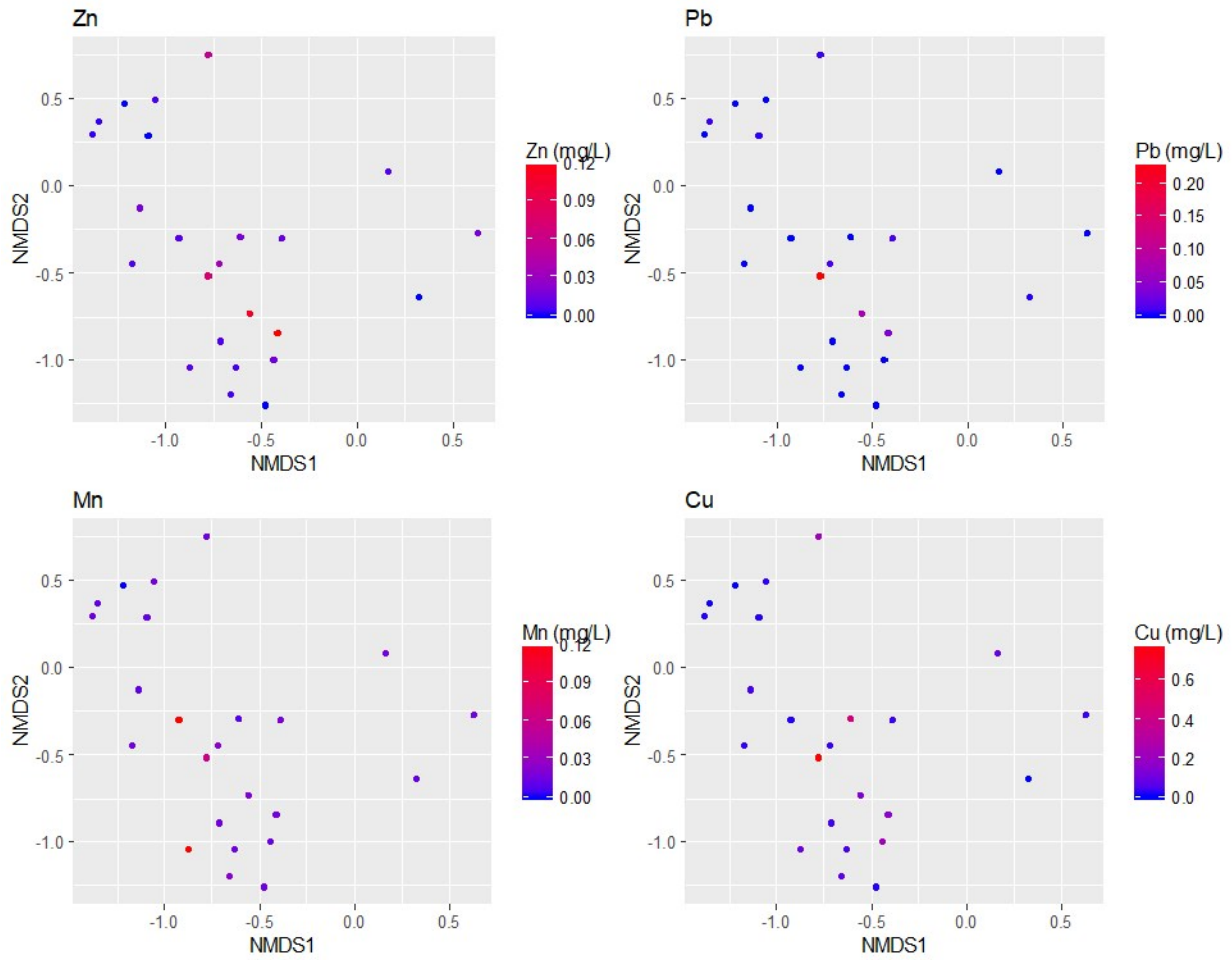
267

268

269

270

271



272

273

274

275

276

277

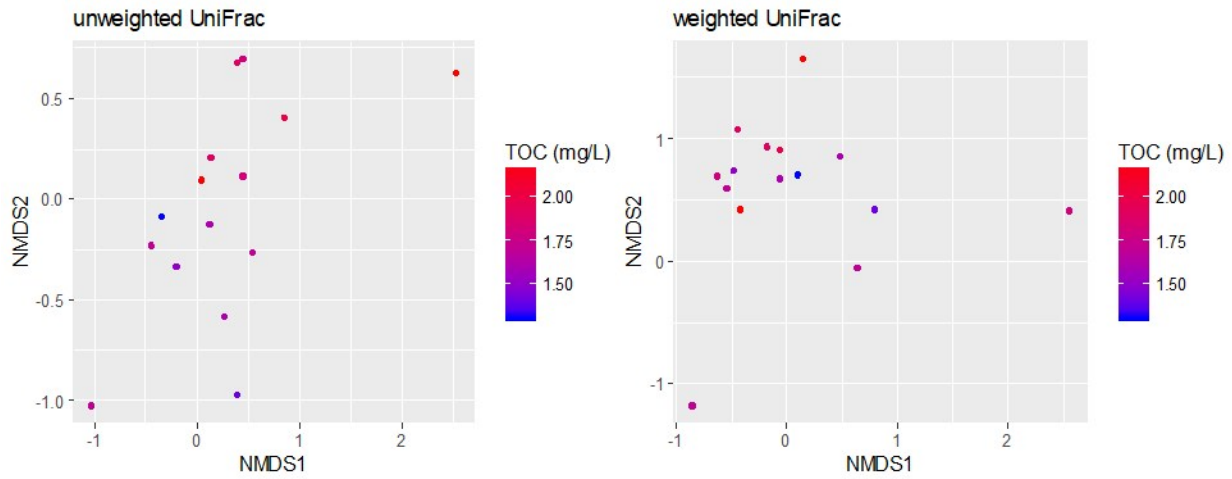
278

279

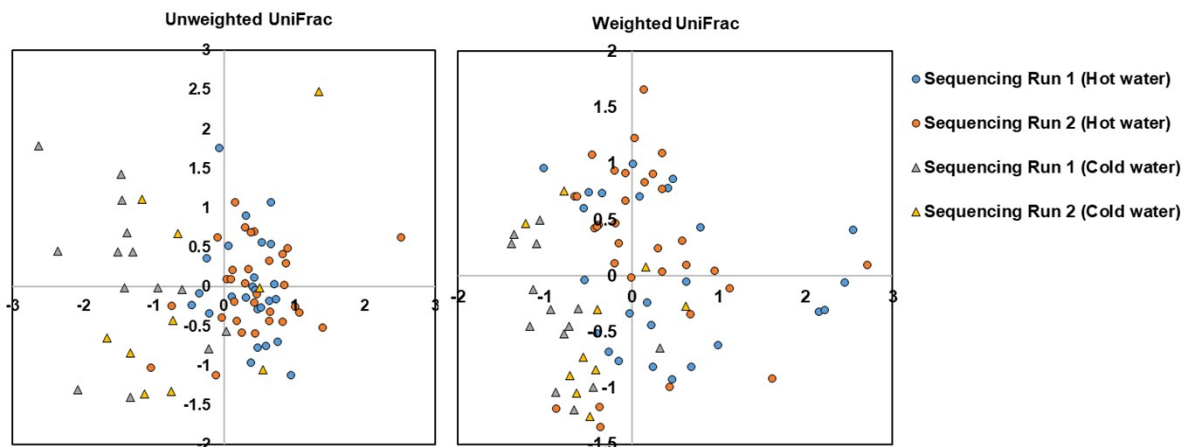
280

281

282



283 SI-Figure 5. Non-metric Multidimensional Scaling (NMDS) plot based on unweighted and  
284 weighted UniFrac distances showing hot water return line bacterial community in relation with  
285 total organic carbon  
286



287

288 SI-Figure 6. Non-metric Multidimensional Scaling (NMDS) plot based on unweighted and  
289 weighted UniFrac distances highlighting samples from two sequencing runs  
290