



This is a repository copy of *Spatial and temporal analogies in microbial communities in natural drinking water biofilms*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/109806/>

Version: Accepted Version

Article:

Douterelo, I., Jackson, M., Solomon, C. et al. (1 more author) (2016) Spatial and temporal analogies in microbial communities in natural drinking water biofilms. *Science of the Total Environment*. ISSN 0048-9697

<https://doi.org/10.1016/j.scitotenv.2016.12.118>

Article available under the terms of the CC-BY-NC-ND licence
(<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

1 **Spatial and temporal analogies in microbial communities in natural drinking water**
2 **biofilms**

3 **I. Douterelo¹ M. Jackson² C. Solomon² and J. Boxall¹**

4 1. Pennine Water Group, Department of Civil and Structural Engineering, Mappin Street,
5 University of Sheffield, Sheffield, S1 3JD, UK.

6 2. Wessex Water, Claverton Down Rd, Bath, Somerset BA2 7WW, UK.

7 Correspondent author: Isabel Douterelo

8 e-mail: i.douterelo@sheffield.ac.uk

9 Telephone: +44 (0)1142 225767

10 Fax: +44 (0) 114 222 5700

11 **Abstract**

12 Biofilms are ubiquitous throughout drinking water distribution systems (DWDS), playing
13 central roles in system performance and delivery of safe clean drinking water. However, little
14 is known about how the interaction of abiotic and biotic factors influence the microbial
15 communities of these biofilms in real systems. Results are presented here from a one-year
16 study using in situ sampling devices installed in two operational systems supplied with
17 different source waters. Independently of the characteristics of the incoming water and
18 marked differences in hydraulic conditions between sites and over time, a core bacterial
19 community was observed in all samples suggesting that internal factors (autogenic) are
20 central in shaping biofilm formation and composition. From this it is apparent that future
21 research and management strategies need to consider the specific microorganisms found to be
22 able to colonise pipe surfaces and form biofilms, such that it might be possible to exclude
23 these and hence protect the supply of safe clean drinking water.

24 **Keywords: bacteria, biofilm, flow, succession, re-growth**

25 **1. Introduction**

26 Drinking Water Distribution Systems (DWDS) are complex ecosystems where biotic and
27 abiotic factors interact along an amalgamate of pipes, storage tanks and other infrastructure
28 extended through vast areas in a buried environment. Many of the interactions are microbially
29 mediated and microorganisms play a central role in determining the quality of the drinking
30 water arriving at customers taps. Most of the microorganisms living in DWDS are attached to
31 pipe surfaces forming mixed-species biofilms. Biofilms can be considered as microbial
32 factories in constant operation where specific processes can take place such as pipe corrosion,
33 residual disinfectant decay or trapping/accumulation of inorganics. Determining the potential

34 for biofilm growth and their composition and structure in DWDS is essential, since biofilm
35 affect the performance of these systems and ultimately the delivery of safe clean drinking
36 water. Understanding the effect of environmental change on biofilm composition and
37 structure in DWDS is challenging mainly due to the difficulty of accessing these buried
38 ecosystems. Commonly observed effects of external factors on biofilms in DWDS are
39 changes in microbial composition and structure (Douterelo et al., 2014), in the components of
40 the extracellular polymeric matrix (EPS) (Fish et al., 2015), changes in density and in
41 chemical and electrical properties (Mukherjee et al., 2012; Janjaroen et al., 2013) and in cell-
42 cell interaction (quorum sensing) (Lee et al., 2014). However, to what extent the combination
43 of biotic and abiotic factors, under realistic conditions, affects the development and
44 composition of natural biofilms in DWDS remains unknown.

45 We have only a limited vision of the microbial ecology of DWDS since most studies have
46 generally focused on free-living organisms from tap samples (e.g. Holinger et al., 2014;
47 Donohue et al., 2015) or water treatment plants (Kasuga et al., 2010; Liu et al., 2013).
48 Studies that have explored biofilms have tended to be in artificial systems in isolation without
49 establishing associations with realistic environmental parameters or between free-living
50 organisms and the attached communities co-habiting the same system (Giao et al., 2008;
51 Moritz et al., 2010). Studies in different environments have shown that natural biofilms
52 respond to their environment and adapt to changes by means of a diverse range of
53 mechanisms (Stewart and Franklin, 2008). Limited information exists regarding whether
54 there are common microbial patterns in biofilms dynamics over time and across different
55 locations. To have a better insight into microbial assembly of natural biofilms and ecological
56 factors influencing their development in DWDS we have used a short-term approach (re-
57 growth every three months) and a long term approach (succession over a one-year period).
58 This will provide a comprehensive vision of how biofilm develop in real systems, allowing
59 for exploration of patterns of behaviour such as seasonal shifts in the structure and
60 composition of biofilms in DWDS.

61 Previous research suggests that the microbial ecology of DWDS will be affected by source
62 water characteristics (Gomez-Alvarez et al., 2015) the type of treatment (Hwang et al., 2012;
63 Pinto et al., 2012) and hydraulic conditions in the system (Douterelo et al., 2013). However,
64 what remains unknown is to which extend external variation will affect attachment to the
65 pipes of certain microorganisms and biofilm formation, composition and dynamics over time.
66 The objectives of the research reported here are to establish the effect of external factors,

67 including different source waters, on the microbial ecology within operational DWDS and to
68 examine patterns of biofilm formation and growth that can inform efficient management of
69 these systems.

70 **2. Materials and Methods**

71 **2.1 Biofilm sampling devices and sampling sites**

72 In situ biofilm sampling devices (Fig. 1) made of High Density Polyethylene (HDPE) were
73 used to study two DWDS with different source waters (physico-chemical characteristics
74 shown in Table 1) and hydraulic regimes (Fig. 2). Each sampling device was fitted into a real
75 system in an available space of 150 mm at both sites and contained 10 modified (increased
76 diameter to improve representative sampling) Pennine Water Group (PWG) coupons (Deines
77 et al., 2010) that enable the study of naturally occurring biofilms in situ without the need for
78 cutting, scraping or flushing the pipes. In addition, the use of PWG coupons allows for
79 studying biofilms on pipe surfaces without distorting boundary layer hydraulic conditions
80 including shear stress and turbulence driven processes such as nutrient exchange. Using these
81 coupons two different processes were studied: 1) quarterly biofilm re-growth and 2) biofilm
82 succession over a one-year period. The sampling devices were first installed in February
83 2013, with first assessment of 3 month-old biofilm development used to test a range of
84 different techniques to evaluate best biofilm monitoring practices (Douterelo et al., 2016).
85 From May 2013 and every three months thereafter the same three coupons were replaced
86 with sterile coupons in order to study biofilm re-growth dynamics at different seasons starting
87 from a completely clean coupon surface (Fig. 1, coupons 2, 5 and 10). Succession coupons
88 (also in triplicate) were sampled at both sites after 3 months (November to February), 6
89 months (August to February) and 12 months (February to February). Bulk water samples
90 were collected whenever site was visited for the above coupon collection. It should be noted
91 that throughout the period there was no other disturbance of the pipe line other than due to
92 hydraulic changes as captured in Fig. 2, such that the biofilm on the surrounding pipe
93 surfaces was not disturbed.

94 One of the sites was supplied with surface water from local springs and river abstraction. The
95 water is treated by coagulation with aluminium sulphate, flocculation and removal of flocs by
96 dissolved air flotation. The water is finally filtered using sand filtration and subsequently
97 granular activated carbon is used to absorb and remove organics. A chlorine disinfectant
98 residual is used in the system. The other site is supplied with groundwater from a standalone
99 borehole site, supplied with a mixture of water from 10 boreholes. The water treatment on

100 site includes marginal chlorination using sodium hypochlorite solution for disinfection
101 residual. Both sites are comprised of similar diameter plastic (Medium and HDPE) pipes in
102 and around the sampling locations.

103 **2.2 Water quality analysis**

104 On the dates of coupon collection, samples from the water that supplied the systems were
105 collected for physico-chemical and microbiological analysis via sampling taps located
106 immediately upstream of the biofilm sampling devices. Temperature and pH were measured
107 in situ using a Hanna portable meter and probe HI 991003 96711(Hanna Instruments,
108 Leighton Buzzard, UK). All the other parameters (see Table 1) were measured by later
109 analysis of discrete water samples by an UK-accredited drinking water laboratory. Flow was
110 measured at 15 minute resolution by magnetic flow meters upstream of the coupon devices.
111 Heterotrophic plate counts - were performed after incubation at 37 °C for 48 h (2-day colony)
112 and 22 °C for 72 h (3-day colony) following UK Standard Methods. For E coli counts a
113 volume of 100 ml of sample was filtered through a 0.45 µm membrane and placed onto the
114 surface of a plate containing membrane lactose glucuronide agar. The plate was then placed
115 in an incubator set to provide pre-incubation for 4 ± 0.25 hours at 30°C followed by an
116 incubation period at 37°C for a minimum of 14 hours. After incubation colonies were
117 counted and the number reported as cfu per 100 ml.

118 **2.3 DNA extraction**

119 DNA was extracted from biofilm (n= 29) and water samples (n=17). For the bulk water
120 samples, three replicates of 2L per site and sampling event were filtered through 0.22-µm
121 nitrocellulose membrane filters (Millipore, Corp.) for subsequent DNA analysis. To extract
122 biofilm from the coupons surface, first the two symmetric outer areas of each coupon were
123 brushed to remove biofilm following the procedure used by Deines et al. (2010). After
124 brushing biofilm suspensions were concentrated in membrane filters as previously explained
125 (Douterelo et al., 2016). DNA was extracted using a method based on proteinase K digestion
126 using a cetyltrimethylammonium bromide (CTAB) protocol followed by further DNA
127 purification using phenol//isoamyl alcohol protocol (Neufeld et al., 2007).

128 **2.4 Sequencing analysis**

129 Sequencing was performed using Illumina Miseq technology with the pair-end protocol by
130 Research and Testing Laboratories (Lubbock, TX, US) using primers 28F
131 GAGTTTGATCNTGGCTCAG and 519 RGTNTTACNGCGGCKGCTG. Paired end reads
132 were merged and denoised via Research and Testing Laboratory Pipeline to remove short

133 sequences, singletons and noisy reads. Chimera were detected using UCHIME (Edgar et al.,
134 2011) and removed from further analysis. Sequences were clustered in Operational
135 Taxonomic Units (OTUs) and selected using UPARSE (Edgar, 2013). Taxonomic
136 assignments with USEARCH global alignment program (Edgar, 2013).

137 An OTUs table at 97% sequence similarity cut off was imported into the software Explicit
138 2.140.5 (Robertson et al., 2013) and a heatmap was created representing the most abundant
139 taxonomic groups with a relative abundance > 0.5%. All the taxonomic groups with less than
140 0.5 % of relative abundance are represented as “Other” in the heatmap. The number of shared
141 OTUs between samples at 97 % sequence similarity cut off and the Venn diagrams were
142 calculated using the web tool provided by the Bioinformatics & Evolutionary Genomics
143 group at the University of Gent <http://bioinformatics.psb.ugent.be/webtools/Venn/>). The
144 software PAST v3.12 (Hammer et al., 2001) was used to estimate Alpha-diversity at 97%
145 sequence similarity and the Shannon diversity index, Chao-I and Dominance-H were
146 calculated. Briefly, the Shannon index (H) measures diversity taking into account the number
147 of OTUs as well as number of taxa, this index varies from 0 for communities with only a
148 single taxon to higher values (max < 5 in this study). Chao 1, is an estimate of total OTUs
149 richness using abundance and occurrence. The Dominance index (1-Simpson index) ranges
150 from 0 (all taxa are equally present) to 1 (one taxon dominates the community completely)
151 (Harper et al., 1999). Approximate confidence intervals for these indexes were computed
152 with a bootstrap procedure (default 9999) and a 95% confidence interval was then calculated.
153 Analysis of similarities (ANOSIM) description was performed based on Bray–Curtis
154 dissimilarity distance matrices to test the differences in community composition among
155 groups of samples using PAST v3.12. The Bray Curtis dissimilarity matrixes were visualised
156 using non-metric multi-dimensional scaling (MDS) diagrams.

157 Sequencing data were deposited in the National Centre for Biotechnology Information
158 (NCBI), Sequence Read Archive SRA SRP095264.

159 **2.4 Statistics**

160 Correlations between physico-chemical and biological parameters were explored by
161 Spearman’s rank non parametric correlations using SPSS 22. Only those parameters showing
162 enough variability between samples were used to establish correlations. Alpha-diversity
163 metrics and the relative abundance of the most abundant OTUs were used as biotic
164 parameters in the establishment of correlations.

165 **3. Results**

166 **3.1 Characteristics of the water supplied to the systems**

167 Figure 2 shows the flows in both systems over the sampled period. Negative or no data on the
168 graph indicate no data collected at that specific time. The surface water (SW) site showed
169 periods of different flow over the period; 0.3-0.9 MI/day during January-June 2013 to a
170 maximum increase in July of 1.7 MI/day and minimum values between the end of July and
171 the end of October 0.1-0.6 MI/day and higher flows November to January 0.3 to 1.4 MI/day.
172 The ground water (GW) flow had less change in flow over the monitored period; up to July
173 the monitoring devices showed flows of less than 0.1 MI/day, after this the average flow
174 increased and was stable ranging typically between 0.1 to 0.35 MI/day. The shifts in flow
175 patterns at both sites were the result of operation changes in the surrounding networks,
176 outside the influence of this study.

177 No colonies counts or E. coli were detected in any of the discrete samples analysed at the
178 time of coupon collection. Temperature ranged from 5.3 °C for SW and 8.5 °C for GW in
179 February 2014 to maximum values of 21.8 °C for SW and 14.1 °C GW in August 2014.
180 Turbidity was stable for GW 0.1 to 0.12 NTU and fluctuated slightly for SW 0.05 in
181 November 2013 to 0.12 in August 2013. pH values were stable for GW 7.5 to 7.7 and slightly
182 higher for SW 7.4 to 8. Other parameters such as conductivity, alkalinity, nitrate and sulphate
183 were higher for the GW samples. Both sites have a similar free chlorine residue, 0.2 to 0.35
184 mg/l, and the levels of chlorine and Total Organic Carbon (TOC) were similar for both sites.

185 **3.2 Seasonality and microbial quarterly dynamics (plankton and biofilms)**

186 **3.2.1 Taxonomy variability: Heatmap**

187 Changes in the relative abundance of different bacterial OTUs were observed between
188 different habitats and over time (seasonality) (Fig. 3). Please note that from two of the
189 samples no sequencing data was obtained these are one planktonic GW sample from
190 November 2013 and one biofilm GW sample from November 2013 . The Heatmap represents
191 the relative OTU abundance of all replicates but in this section to summarise the information
192 the average of replicates was calculated according to habitat and season. The bacterial
193 community of both biofilm and planktonic samples was dominated by Proteobacteria with
194 average relative abundance of up to 79 %, (Fig. 3). This phylum was followed by
195 Actinobacteria (8.3%) and Firmicutes (5 %) that was commonly found in SW planktonic
196 communities. Within the Alphaproteobacteria the order Sphingomonadales accounted for
197 16.6% of the sequences recovered from all samples and the genus Sphingomonas (4.3%) was
198 present in all biofilm samples and in SW planktonic samples. The genus Hypomicrobium

199 (4.7%) was particularly abundant in SW and GW samples in August and February and SW
200 biofilms in August.

201 SW planktonic communities showed clear seasonal changes, August samples were dominated
202 by Alphaproteobacteria and the genus *Hypomicrobium* (> 24%) while in November
203 Gammaproteobacteria (> 30%) was the main represented phylum in the samples with more
204 than 14% of those affiliated to the Pseudomonadales order. Actinobacteria were also
205 abundant (>23%) in November with the genus *Arthobacter* (>7%) highly present in the
206 samples. In February there was an increase again in Alphaproteobacteria (39 %) in the total
207 community mainly represented by the order Brucellaceae (16%) and with 21% of
208 Actinobacteria with *Rhodococcus* as the main genus (9%).

209 GW planktonic communities showed less variability over time than SW but the relative
210 abundance of different taxonomic groups also changed. Alphaproteobacteria was the main
211 represented phylum in all the seasons ranging from 38% in August to 25% in February.
212 Within the Alphaproteobacteria phylum the order Sphingomonales dominated in August
213 (30%) and November (45%). In August Firmicutes was highly represented by the order
214 Clostridia (18%), however in November and February this group has a minor representation
215 in the total bacteria community. Gammaproteobacteria was presented in all months but was
216 particularly abundant in February (25%) being the main order the Pseudomonales (16%).
217 During this month Actinobacteria (23%) was an important component of the microbial
218 community with the genus *Mycobacterium* (6%) as the main representative.

219 Biofilm samples maintained several dominant OTUs over time mainly *Pseudomonas* and a
220 high quantity of low abundance (less than 0.5% relative abundance) or rare OTUs. In SW
221 biofilm samples Alphaproteobacteria was highly represented (>73% in August 2013), mainly
222 by genera such as *Brucella* (3%), *Hypomicrobium* (6%), *Sphingomonas* (13%) and
223 *Sphingopsis* (34%). Gammaproteobacteria was commonly found in all seasons but
224 particularly in November (67%) and February (57%) with *Pseudomonas* as the main genus.
225 For GW biofilms *Pseudomonas* was markedly abundant in all the seasons with an average
226 representation always higher than 45 %. Alphaproteobacteria abundance changed over time
227 from 34% in August 2013 to 13% in February 2014. In August 2013 *Brevundimonas* was
228 representing 20% of the total community but its abundance decreased to only 3% in
229 February. *Sphingomonas* was highly abundant in all seasons but mainly in February (7%). In
230 all the biofilm samples the main differences in microbial community structure over time are

231 determined by several OTUs with less than 0.5 % relative abundance, marked in the heatmap
232 as “Others”.

233 **3.2.2 Shared OTUs and MDS**

234 Figure 4A shows Venn diagrams for each sampling season with the number of unique and
235 shared OTUs between bacterial communities. In August, all habitats shared a large number of
236 OTUs 87 between biofilm samples and 71 between planktonic communities. Conversely, a
237 very low number of OTUs were shared in autumn, where SW and GW biofilm samples only
238 shared 7 OTUs. The MDS analysis (Fig. 4B) using Bray-Curtis dissimilarity matrix (at 97%
239 sequence similarity cut off) showed high variability between planktonic samples and a more
240 stable community over time for biofilm samples from both SW and GW. Bacterial
241 community structure was highly variable for planktonic SW samples. Less variability
242 between biofilm samples was observed indicating that bacterial assemblages tend to re-
243 growth in the same way independently of the water source supplied and the operational
244 parameters. Differences in the community composition between samples were tested using
245 ANOSIM (Fig. 4C). However, no statistical significant differences were observed over time
246 (seasonality) between samples. Statistical differences were observed between habitats
247 (locations), ANOSIM showed significant differences between planktonic samples ($p < 0.01$)
248 from SW and GW. Biofilm and planktonic samples also showed significant difference for
249 each sampling site SW ($p < 0.01$) and GW ($p < 0.01$). No significant differences were obtained
250 for biofilm samples at both sites (SW vs. GW).

251 **3.2.3 Alpha-diversity: diversity, richness and dominance**

252 The alpha-diversity metrics of bacterial communities (Fig. 5) indicate clear changes between
253 habitats and seasons. In general, Shannon diversity index (Fig. 5A) exhibited higher values
254 for planktonic communities than for the biofilm ones. SW planktonic samples showed higher
255 diversity when compared with GW. For biofilm samples, diversity was higher for SW
256 biofilms in August 2013 but less in November 2013 and similar levels were showed for
257 February 2014 where all biofilm samples showed an increase in diversity. The high diversity
258 in February 2014 in the planktonic samples does not correspond with the low diversity in SW
259 biofilms for that time.

260 Chao richness (Fig. 5B) for planktonic communities was higher for SW than GW sample,
261 with the exception of November 2013 samples, and similar for biofilm samples from both
262 sites. Dominance (Fig. 5C) was low for all water samples (< 0.3) but high for biofilms with

263 most of the samples presenting values above 0.3 indicating the dominance of fewer OTUs in
264 these communities.

265 **3.3 Biofilm community composition and structure succession analysis (long term one** 266 **year experiment)**

267 **3.3.1 Taxonomic variability**

268 All the samples (Fig. 6) showed high presence of Gammaproteobacteria, mainly the genus
269 Pseudomonas, particularly the samples experiencing only 3 months of biofilm development.
270 Sphingomonadales and Cyanobacteria were consistently abundant in all the samples.
271 Pseudomonas dominated in SW biofilm samples during the 6 months of biofilm
272 development. Other initially abundant OTUs decreased with time, such as Rhizobiales
273 (Alphaproteobacteria) from 8% to 2.5% and Acidovorax (Betaproteobacteria) from 13% in 3
274 month-old biofilms to 0.7 %. However, several OTUs increased over time and in 12 month-
275 old SW biofilm samples a more diverse community was present with high abundance of
276 Alphaproteobacteria. Alphaproteobacteria increased over time from 14 to 28% (e.g.
277 Sphingomonas). Firmicutes increased from 0.5% to 15% and Actinobacteria from 0.9% to
278 13% with the main representative genera being Mycobacterium and Propionibacterium.

279 The GW community consistently showed predominance of Pseudomonas (> 35%) and to a
280 certain extent of Sphingomonas (1.5% to 14%). Biofilm samples from 6 and 12-month
281 exposure displayed high relative abundance of Mycobacterium (0.5-55%) and
282 Brevundimonas (3-39%).

283 **3.3.2 Shared OTUs and MDS**

284 In SW biofilm samples, the proportion of unique vs. shared OTUs was similar over time (Fig.
285 7A). GW biofilm communities showed less temporal changes and the number of OTUs
286 shared decreased between 6 and 12 month-old samples. For GW the number of shared OTUs
287 decreased and between 3 and 12 month-old samples (Fig 7B). The number of OTUs unique to
288 GW biofilm samples on 12 month-old samples when compared to 3 month-old samples was
289 74. SW samples displayed a slightly increase in the number of OTUs in more mature biofilms
290 but the number of unique OTUs for GW 3 month-old samples decreased. Despite SW and
291 GW samples being more different to start with (3 month-old biofilms), regardless of sharing a
292 relatively high percentage of OTUs, more mature biofilms were less distinctive and GW
293 samples had only 37 unique OTUs when compared with SW samples. The number of OTUs
294 unique to GW 12 month-old samples when compared to SW 12 month-old samples was only
295 37.

296 The compositional comparison of samples in a non-metric multidimensional scaling (MDS)
297 plot (Fig. 7C) showed difference over time among sites and high variability for SW
298 replicates. Temporal changes were more marked for SW samples, those showed high
299 variability between samples. The 3 month old SW samples and the GW samples from 3, 6
300 and 12 month-old biofilms showed a certain degree of clustering, while the SW 6 and 12
301 month-old samples were more scattered and dissimilar. The ANOSIM analysis showed that
302 when each location was analysed independently non-significant statistical differences
303 between months were observed (Fig. 7D). However, analysis of all samples for each habitat
304 showed significant differences ($R= 0.122$ and $p =0.0189$) between SW and GW samples

305 **3.3.3 Alpha-diversity: diversity, richness and dominance**

306 Shannon diversity index presented similar levels for SW and GW biofilms for the 6 month-
307 old biofilm samples and increased for 12 month-old samples particularly for SW (Fig. 8A). In
308 the GW, the Shannon index was on average slightly lower in the 3 month-old samples, but
309 was similar for the 6 and 12 month-old biofilm samples. Chao richness indicator (Fig. 8B)
310 was higher for SW samples than for GW ones. In the SW samples the richness tended to
311 increase from 3 to 6 month-old biofilms and then decreased for the 12 month-old samples. In
312 GW, the Chao richness estimator was low for 3 month-old biofilm and then tended to
313 increases from 3 to 6 to 12 month-old samples. In general, the dominance indicator (Fig. 8C)
314 was slightly higher for SW samples for 3 and 6 month-old biofilms compared to GW, but
315 decreased notably for the 12 month-old SW samples. GW samples had on average similar
316 dominance levels between 3, 6 and 12 month-old biofilms.

317 **3.4 Relation of physico-chemistry with microbiological parameters**

318 Spearman's rank correlations were calculated to explore the correlation between different
319 physico-chemical and microbiological parameters. These are shown in supplementary
320 material.

321 Flow was strongly positively correlated with pH and sulphate but negatively correlated with
322 conductivity, alkalinity and nitrate. Flow was positively correlated with *Sphingomonas*,
323 *Nevskia*, *Brucella*, *Rhodococcus* and *Propionibacterium* and negatively with *Mycobacterium*.
324 Significant positive correlations were detected between temperature and conductivity,
325 alkalinity, TOC and pH. There was no correlation with chlorine, but levels were similar
326 between sites and over time. There were also correlations with Chao and the relative
327 abundance of several bacterial genera, including *Brevundimonas*, *Hypomicrobium*,
328 *Erythrobacter* and negatively with *Pseudomonas* and *Rhodanobacter*. Turbidity was

329 significantly positive correlated with levels of TOC, pH, Cl and the abundance of
330 *Rhodococcus*, *Brucella* and *Hypomicrobium* among others. Conductivity and alkalinity and
331 nitrate were also positively correlated.

332 Correlations between specific bacterial genera were also found. For example, the main OTU
333 presented in biofilm samples, *Pseudomonas*, correlated positively with *Sphingomonas* and
334 *Acidovorax* and was negatively correlated with *Propionibacterium*, *Staphylococcus* and
335 *Erythrobacter*. *Nevskia* and *Sphingopyxis* were correlated with most of the other highly
336 abundant OTUs in biofilms including *Rhodococcus*, *Propionibacterium*, *Brevundimonas*,
337 *Erythrobacter* and *Sphingomonas*.

338 **4. Discussion**

339 Differences were seen in the taxonomic composition of SW and GW supplied systems,
340 particularly in planktonic communities, with GW samples showing less marked changes over
341 time when compared with SW. The variability in seasonal external factors had an effect
342 particularly on SW planktonic communities. This was expected and similar observations were
343 made in previous studies such as Gomez Alvarez et al (2015), which showed differences
344 between GW and SW supplied drinking water-related systems in planktonic communities.
345 Planktonic communities are commonly considered as the source of bacteria colonisers in
346 DWDS biofilms (Henne et al., 2012). However, this study shows significant differences in
347 the bacterial composition of these habitats (water vs. biofilm) and a lack of significant
348 statistical influence of changes in the free-living organisms on the biofilm structure.

349 Ling et al (2016) showed that seasonality was the main contributor to community structure
350 variation by studying biofilm samples from household water meters and in tap water samples.
351 Similarly, Bachmann and Edyvean (2005) reported that the origin of raw water had a great
352 impact on the bacterial communities in DWDS and Pinto et al (2014) found that the
353 planktonic community was strongly correlated with the community found in water treatment
354 works filters. However, unlikely these previous studies successional changes were detected in
355 biofilm communities in the long term 1 year experiment (Fig. 7C) but no clear seasonal effect
356 was observed on the quarterly samples (Fig 4B). Consequently, changes in the bulk water
357 communities did not exert a clear influence on the composition of the attached community.

358 A major result of this study is that the biofilm communities from both quarterly re-growth
359 and succession over a year presented a similar core microbial community between both sites,
360 and distinct from the planktonic community (as shown in Fig. 4B and 7C). The importance of
361 external factors such as flow rate (Lehtola et al., 2006; Manuel et al., 2007), chlorine

362 (Butterfield et al., 2002; Ndiongue et al., 2005), nutrient supply (Chandy and Angles, 2001;
363 Boe-Hansen et al., 2003) and pipe material (Niquette et al., 2000) on biofilm development
364 has been highlighted in other studies under controlled laboratory conditions. Here where
365 biofilms were grown under natural, non-manipulated conditions, the main factors correlating
366 with the distribution of certain bacterial groups were flow rate, TOC, temperature, pH and
367 sulphate. The average flow rate was correlated with the relative abundance of several core
368 bacteria consequently the presence or absence of some of these bacteria can be associated
369 with the hydraulic conditions in the system. Previous work suggested that hydraulic
370 conditions and shear forces have an influence on DWDS biofilms (Douterelo et al 2013,
371 2016). Douterelo et al (2016) showed such a difference across biofilm amount, strength and
372 community composition for the same two sites as studied further here. However, it should be
373 noted that the period covered in Douterelo et al (2016) was exclusively from the near
374 stagnant flow period at the GW site (Fig. 2). All 3 and 6 month-old samples reported here are
375 from the flowing period of the GW site, and the 12 month-old samples were dominated by
376 the flowing condition at this site. This suggests that while very low flow may lead to a
377 different community and more and weaker biofilm material, flow rate might not be a central
378 factor in shaping the dominant members of biofilm communities once a sustained regularly
379 (daily) occurring turbulent flow regime is experienced. Douterelo et al (2013) and (2014),
380 both using a fully representative plastic pipe system under laboratory conditions, observed
381 the influence of hydraulic regimes on biofilm structure, physical strength and discolouration
382 risk. However, while representative of real system the conditions of these studies were
383 controlled and repeated, hence it is possible that the unavoidable variations in daily
384 conditions for the operational system studied here over such a long study period both at and
385 between the sites obfuscated any such effects.

386 Based on the outcomes of this research, where a consistent core biofilm community was
387 found independently of the sampling location, we can conclude that a group of bacteria that
388 are adapted to DWDS are ubiquitous in these systems. We therefore hypothesise that there
389 are internal factors related with the composition of the biofilm per se that are shaping the
390 diversity of biofilms. The concept of a core community forming part of biofilms was first
391 observed by Henne et al (2011) using a molecular fingerprinting technique known as Single
392 Strand Conformation Polymorphism (SSCP) in a non-chlorinated distribution system in
393 Germany and has been corroborated by Ling et al (2016) in water meters biofilms using
394 pyrosequencing. The present study reinforces the existence of a universal core community of

395 microorganisms in DWDS biofilm by using a high-throughput sequencing method and in
396 chlorinated systems supplied with different water sources and hydraulic regimes. The
397 methodology used here involved the insertion of relatively small areas of sterile, autoclaved,
398 pipe surface within a larger system that was not otherwise disturbed or impacted. This taken
399 with the lack of significant influence of the planktonic community on the biofilm structure
400 leads to the suggestion that this core community was predominately influenced by the
401 surrounding biofilm composition from the local or upstream pipes. It is interesting to note
402 that most likely internal regulatory factors were dominating rapidly such that the majority of
403 the biofilm communities change little over the 1 year of succession, although dominance does
404 drop for the SW 12 month-old samples (Fig. 8C). This is in agreement with Lyautey et al.
405 (2005) that if microbial succession is the predominant mechanism of temporal changes in
406 community structure, then these changes should be repeatable and predictable for a given
407 region. If the temporal changes observed in biofilm communities were autogenic this should
408 have led to comparable communities over the seasons, as seen here. It is interesting to
409 compare this observation with the findings of Douterelo et al (2016) where biofilm
410 community was shown to evolve over time in response to repeated flushing of an operational
411 system, the flushing was observed to exerting a selective pressure on the biofilms, and that
412 microbial dynamics were influenced by changes in water source parameters particularly
413 phosphate and metals. However, flushing does not remove all material and the remaining
414 biofilm left attached on the pipe walls will influence the regrowth of new biofilm. Conversely
415 this study, where sterile coupons were reinserted every three months, exhibits a rapidly re-
416 established community from the stable, undisturbed biofilm community of the surrounding
417 pipe surfaces. This again reinforces the influence of the local biofilm community rather than
418 the planktonic community or any other external selective pressure.

419 Biofilm communities studied here were specifically enriched with certain bacteria,
420 predominantly *Pseudomonas*. *Pseudomonas* was the main bacteria forming what can be
421 considered the dominant community of biofilm structures, likely sourced from the
422 surrounding undisturbed biofilm and hence independently of the water source studied. It is
423 known that mixed-species biofilm formation can depend on the presence of species with high
424 affinity and adherence to surfaces such as *Pseudomonas* facilitating the attachment of other
425 microorganisms (Dunne, 2002; Kostakioti et al., 2013). *Pseudomonas* species can easily
426 produce exopolysaccharides (Ghafoor et al., 2011; Irie et al., 2012) that can provide biofilms
427 with a “stabilising effect”. Thus the dominance of *Pseudomonas* independently of any

428 external factor confirms that autogenic factors were shaping the biofilm community
429 composition. The interaction of microorganisms in drinking water-related biofilms has been
430 studied by selecting species isolated from drinking water systems and using dual
431 combinations of them (Simões et al., 2007; Ramalingam et al., 2013). What remains
432 unknown is the understanding of how natural mixed-species biofilm work and interact when
433 there are also a combination of external factors that influence these communities. Here we
434 confirm that in natural DWDS biofilms there is a clear tendency for particular bacteria to
435 positively interact and form biofilms and this happens independently of external factors.
436 These specific interactions can be exploited to exclude undesirable pathogens from healthy
437 biofilm communities and to favour beneficial phenotypes. For example, in this study
438 *Pseudomonas* was negatively correlated with the relative abundance of several bacteria
439 including some potential pathogenic genera such as *Staphylococcus* and *Brucella*.

440 It is clear from looking at the 3, 6 and 12 month-old biofilms that once the core community
441 has been establish changes over time are due to low abundance (rare) bacteria. Minor
442 bacterial representatives (less than 0.5% relative abundance) were those that shaped the
443 overall diversity over time and between sites. Similarly, Holinger et al. (2014) studying the
444 bacteriological composition of tap water samples from different North American cities
445 observed considerable variation among the rare phylotypes and that the most abundant taxa
446 were similar from system to system, regardless of source water type. The author suggested
447 that the similarity among the abundant taxa between systems was the consequence of the
448 selective influence of chlorine-based disinfection and the local environment of the DWDS. In
449 the present study chlorine was correlated with the distribution of a limited number of OTUs
450 but was not the main factor affecting their distribution. However, what remains unknown is
451 the relevance of these diverse minor representatives in the overall function of the system and
452 whether if they play a central role in covering specific functions within biofilms that might
453 change depending on the environment to adapt to different conditions.

454 Overall these experiments show that autogenic factors are important in shaping biofilm
455 composition. In particular, it is surrounding biofilm community that is key in the processes of
456 biofilm development, rather than the incoming bulk water community. Although it should be
457 noted that only plastic pipes where studied here. This means that by simply adjusting or
458 managing the physico-chemical characteristics of the water incoming the system it is not
459 possible to fully control biofilm formation in DWDS. Most biofilm-related research focuses
460 on how to eliminate biofilms but future control strategies might be better based on engineer

461 biofilms to perform specific functions. Further consideration should be given to specific
462 microorganisms able to colonise pipe surfaces and form biofilms, such that it might be
463 possible to exclude adverse free-living organisms from colonising pipes protecting the supply
464 of safe clean drinking water.

465 **Conclusion**

466 The dynamics of drinking water distribution systems biofilms were assessed over a one-year
467 period and biofilm re-growth was analysed at quarterly intervals using sampling devices
468 installed in situ in two operational networks. The two systems were supplied from surface
469 and ground water sources, with only plastic pipework in and around the sampling locations.
470 The planktonic communities clearly changed over time, particularly those supplied with
471 surface water, influenced by seasonal changes. Independently of these and other
472 characteristics of the incoming water between sites, the biofilm communities shared a high
473 number of common and highly abundant operational taxonomic units (OTUs). The genus
474 *Pseudomonas* was the main inhabitant of the biofilms independently of the network studied,
475 forming part of a dominant core community ubiquitous to all biofilm samples, irrespective of
476 if they were from succession over the one-year period or re-growth samples from quarterly
477 intervals. The less abundant bacterial representatives (OTUs <0.5 % relative abundance), rare
478 OTUs, were responsible for most of the variability over time and between habitats. The main
479 abiotic factors affecting the microbiology of the systems included flow rate, temperature and
480 pH.

481 This research shows that while abiotic factors may influence the amount and strength of
482 biofilm, its composition was strongly influenced by the biofilm community already present
483 within the surrounding pipes. Such detailed understanding of the process of biofilm formation
484 in DWDS is essential for the management of these systems for the delivery of safe clean
485 drinking water, such as the potential to control or manipulate certain key bacteria to limit
486 formation.

487 **Acknowledgments**

488 The research reported here was supported by the UK Engineering and Physical Sciences
489 Research Council (EPSRC) Challenging Engineering grant EP/G029946/1, the EPSRC
490 Platform Grant Funding EP/1029346/1 and the EPSRC-LWEC Challenge Fellowship
491 EP/N02950X/1. We would like to thank Wessex Water Services Ltd for sampling sites, field
492 work and physico-chemical sample analysis.

493

494 **REFERENCES**

- 495 Bachmann, R.T., Edyvean, R.G.J., 2005. Biofouling: an historic and contemporary review of
496 its causes, consequences and control in drinking water distribution systems. *Biofilms* 2, 197-
497 227.
- 498 Boe-Hansen, R., Martiny, A.C., Arvin, E., Albrechtsen, H.J., 2003. Monitoring biofilm
499 formation and activity in drinking water distribution networks under oligotrophic conditions.
500 *Water science and technology : a journal of the International Association on Water Pollution*
501 *Research* 47, 91-97.
- 502 Butterfield, P.W., Camper, A.K., Ellis, B.D., Jones, W.L., 2002. Chlorination of model
503 drinking water biofilm: implications for growth and organic carbon removal. *Water Res* 36,
504 4391-4405.
- 505 Chandy, J.P., Angles, M.L., 2001. Determination of nutrients limiting biofilm formation and
506 the subsequent impact on disinfectant decay. *Water Res* 35, 2677-2682.
- 507 Deines, P., Sekar, R., Husband, P.S., Boxall, J.B., Osborn, A.M., Biggs, C.A., 2010. A new
508 coupon design for simultaneous analysis of in situ microbial biofilm formation and
509 community structure in drinking water distribution systems. *Appl Microbiol Biot* 87, 749-
510 756.
- 511 Donohue, M.J., Mistry, J.H., Donohue, J.M., O'Connell, K., King, D., Byran, J., Covert, T.,
512 Pfaller, S., 2015. Increased Frequency of Nontuberculous Mycobacteria Detection at Potable
513 Water Taps within the United States. *Environmental Science & Technology* 49, 6127-6133.
- 514 Douterelo, I., Husband, S., Boxall, J.B., 2014. The bacteriological composition of biomass
515 recovered by flushing an operational drinking water distribution system. *Water Res* 54, 100-
516 114.
- 517 Douterelo, I., Jackson, M., Solomon, C., Boxall, J., 2016. Microbial analysis of in situ
518 biofilm formation in drinking water distribution systems: implications for monitoring and
519 control of drinking water quality. *Appl Microbiol Biot* 100, 3301-3311.
- 520 Douterelo, I., Sharpe, R.L., Boxall, J.B., 2013. Influence of hydraulic regimes on bacterial
521 community structure and composition in an experimental drinking water distribution system.
522 *Water Res* 47, 503-516.
- 523 Dunne, W.M., 2002. Bacterial adhesion: Seen any good biofilms lately? *Clin Microbiol Rev*
524 15, 155-+.
- 525 Edgar, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon
526 reads. *Nat Methods* 10, 996-998.
- 527 Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R., 2011. UCHIME improves
528 sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194-2200.
- 529 Fish, K.E., Collins, R., Green, N.H., Sharpe, R.L., Douterelo, I., Osborn, A.M., Boxall, J.B.,
530 2015. Characterisation of the Physical Composition and Microbial Community Structure of
531 Biofilms within a Model Full-Scale Drinking Water Distribution System. *PLoS ONE* 10,
532 e0115824.
- 533 Ghafoor, A., Hay, I.D., Rehm, B.H.A., 2011. Role of Exopolysaccharides in *Pseudomonas*
534 *aeruginosa* Biofilm Formation and Architecture. *Appl Environ Microb* 77, 5238-5246.

535 Giao, M.S., Azevedo, N.F., Wilks, S.A., Vieira, M.J., Keevil, C.W., 2008. Persistence of
536 *Helicobacter pylori* in heterotrophic drinking-water biofilms. *Appl Environ Microb* 74, 5898-
537 5904.

538 Gomez-Alvarez, V., Humrighouse, B.W., Revetta, R.P., Santo Domingo, J.W., 2015.
539 Bacterial composition in a metropolitan drinking water distribution system utilizing different
540 source waters. *Journal of Water and Health* 13, 140-151.

541 Hammer, Ø. & Harper, D.A.T. 2006. *Paleontological Data Analysis*. Blackwell.

542 Hammer, Ø., Harper, D.A.T., and P. D. Ryan, 2001. PAST: Paleontological Statistics
543 Software Package for Education and Data Analysis. *Palaeontologia Electronica* 4(1): 9pp.

544 Harper, D.A.T. (ed.). 1999. *Numerical Palaeobiology*. John Wiley & Sons.

545 Henne, K., Kahlisch, L., Brettar, I., Höfle, M.G., 2012. Analysis of Structure and
546 Composition of Bacterial Core Communities in Mature Drinking Water Biofilms and Bulk
547 Water of a Citywide Network in Germany. *Appl Environ Microb* 78, 3530-3538.

548 Holinger, E.P., Ross, K.A., Robertson, C.E., Stevens, M.J., Harris, J.K., Pace, N.R., 2014.
549 Molecular analysis of point-of-use municipal drinking water microbiology. *Water Research*
550 49, 225-235.

551 Hwang, C., Ling, F., Andersen, G.L., LeChevallier, M.W., Liu, W.-T., 2012. Microbial
552 Community Dynamics of an Urban Drinking Water Distribution System Subjected to Phases
553 of Chloramination and Chlorination Treatments. *Applied and Environmental Microbiology*
554 78, 7856-7865.

555 Irie, Y., Borlee, B.R., O'Connor, J.R., Hill, P.J., Harwood, C.S., Wozniak, D.J., Parsek,
556 M.R., 2012. Self-produced exopolysaccharide is a signal that stimulates biofilm formation in
557 *Pseudomonas aeruginosa*. *Proceedings of the National Academy of Sciences* 109, 20632-
558 20636.

559 Janjaroen, D., Ling, F., Monroy, G., Derlon, N., Mogenroth, E., Boppart, S.A., Liu, W.-T.,
560 Nguyen, T.H., 2013. Roles of ionic strength and biofilm roughness on adhesion kinetics of
561 *Escherichia coli* onto groundwater biofilm grown on PVC surfaces. *Water Res* 47, 2531-
562 2542.

563 Kasuga, I., Nakagaki, H., Kurisu, F., Furumai, H., 2010. Predominance of ammonia-
564 oxidizing archaea on granular activated carbon used in a full-scale advanced drinking water
565 treatment plant. *Water Research* 44, 5039-5049.

566 Kostakioti, M., Hadjifrangiskou, M., Hultgren, S.J., 2013. Bacterial Biofilms: Development,
567 Dispersal, and Therapeutic Strategies in the Dawn of the Postantibiotic Era. *Cold Spring*
568 *Harbor Perspectives in Medicine* 3, a010306.

569 Lee, O.O., Wang, Y., Tian, R., Zhang, W., Shek, C.S., Bougouffa, S., Al-Suwailem, A.,
570 Batang, Z.B., Xu, W., Wang, G.C., Zhang, X., Lafi, F.F., Bajic, V.B., Qian, P.Y., 2014. In
571 situ environment rather than substrate type dictates microbial community structure of
572 biofilms in a cold seep system. *Sci Rep* 4.

573 Lehtola, M.J., Laxander, M., Miettinen, I.T., Hirvonen, A., Vartiainen, T., Martikainen, P.J.,
574 2006. The effects of changing water flow velocity on the formation of biofilms and water
575 quality in pilot distribution system consisting of copper or polyethylene pipes. *Water Res* 40,
576 2151-2160.

577 Ling, F., Hwang, C., LeChevallier, M.W., Andersen, G.L., Liu, W.-T., 2016. Core-satellite
578 populations and seasonality of water meter biofilms in a metropolitan drinking water
579 distribution system. *ISME J* 10, 582-595.

580 Liu, G., Ling, F.Q., Magic-Knezev, A., Liu, W.T., Verberk, J.Q.J.C., Van Dijk, J.C., 2013.
581 Quantification and identification of particle-associated bacteria in unchlorinated drinking
582 water from three treatment plants by cultivation-independent methods. *Water Research* 47,
583 3523-3533.

584 Lyautey, E., Jackson, C.R., Cayrou, J., Rols, J.-L., Garabétian, F., 2005. Bacterial
585 Community Succession in Natural River Biofilm Assemblages. *Microbial Ecology* 50, 589-
586 601.

587 Manuel, C.M., Nunes, O.C., Melo, L.F., 2007. Dynamics of drinking water biofilm in
588 flow/non-flow conditions. *Water Res* 41, 551-562.

589 Moritz, M.M., Flemming, H.-C., Wingender, J., 2010. Integration of *Pseudomonas*
590 *aeruginosa* and *Legionella pneumophila* in drinking water biofilms grown on domestic
591 plumbing materials. *International Journal of Hygiene and Environmental Health* 213, 190-
592 197.

593 Mukherjee, J., Karunakaran, E., Biggs, C.A., 2012. Using a multi-faceted approach to
594 determine the changes in bacterial cell surface properties influenced by a biofilm lifestyle.
595 *Biofouling* 28, 1-14.

596 Ndiongue, S., Huck, P.M., Slawson, R.M., 2005. Effects of temperature and biodegradable
597 organic matter on control of biofilms by free chlorine in a model drinking water distribution
598 system. *Water Res* 39, 953-964.

599 Neufeld, J.D., Vohra, J., Dumont, M.G., Lueders, T., Manefield, M., Friedrich, M.W.,
600 Murrell, J.C., 2007. DNA stable-isotope probing. *Nat Protoc* 2, 860-866.

601 Niquette, P., Servais, P., Savoie, R., 2000. Impacts of pipe materials on densities of fixed
602 bacterial biomass in a drinking water distribution system. *Water Res* 34, 1952-1956.

603 Pinto, A.J., Schroeder, J., Lunn, M., Sloan, W., Raskin, L., 2014. Spatial-Temporal Survey
604 and Occupancy-Abundance Modeling To Predict Bacterial Community Dynamics in the
605 Drinking Water Microbiome. *Mbio* 5.

606 Pinto, A.J., Xi, C., Raskin, L., 2012. Bacterial Community Structure in the Drinking Water
607 Microbiome Is Governed by Filtration Processes. *Environmental Science & Technology* 46,
608 8851-8859.

609 Ramalingam, B., Sekar, R., Boxall, J., Biggs, C., 2013. Aggregation and biofilm formation of
610 bacteria isolated from domestic drinking water. *Water Science & Technology: Water Supply*
611 13.

612 Robertson, C.E., Harris, J.K., Wagner, B.D., Granger, D., Browne, K., Tatem, B., Feazel,
613 L.M., Park, K., Pace, N.R., Frank, D.N., 2013. Explicit: graphical user interface software for
614 metadata-driven management, analysis and visualization of microbiome data. *Bioinformatics*
615 29, 3100-3101.

616 Simões, L.C., Simões, M., Vieira, M.J., 2007. Biofilm Interactions between Distinct Bacterial
617 Genera Isolated from Drinking Water. *Applied and Environmental Microbiology* 73, 6192-
618 6200.

619 Stewart, P.S., Franklin, M.J., 2008. Physiological heterogeneity in biofilms. *Nat Rev Micro* 6,
620 199-210.

621 **Table and Figures**

622 **Table 1:** Physico-chemical and microbiological parameters measured during every sampling
623 event in August, November and February.

624 **Figure 1:** **A)** Replacement of a section of pipe for installation of coupon devices **B)** insertion
625 of a coupon in one of the holders of the biofilm sampling device **C)** Coupons after sampling
626 showing areas used for biofilm removal **D)** picture showing the distribution of the coupons in
627 the device, coupons in red (2, 5 and 10) were used for re-growth studies, the other coupons
628 were used for long term studies to study succession.

629 **Figure 2:** Flow regimes for the ground water and surface water sampling site.

630 **Figure 3:** Heatmap showing the taxonomic distribution of quarterly water and biofilm
631 samples studying biofilm re-growth with relative abundance > 0.5%. * indicates that only
632 information from two samples was obtained. **Others** indicate the abundance of bacterial
633 groups with < 0.5% of relative abundance.

634 **Figure 4:** **A)** Venn diagram of the planktonic and biofilm community showing the shared and
635 unique OTUS for samples collected quarterly. **B)** non-parametric multidimensional analysis
636 (MDS) of biofilm and water samples. **C)** table showing the results from the ANOSIM
637 analysis. Labels: **BSW:** Biofilm Surface Water, **BGW:** Biofilm Groundwater, **WSW:** Water
638 Surface Water, **WGW:** Water Ground Water, **Ag:** august, **Nov:** November, **Feb:** February.

639 **Figure 5:** Alpha diversity metrics for biofilm and water samples used to study biofilm re-
640 growth. **A)** Shannon diversity index. **B)** Chao richness indicator. **C)** Dominance-H. Labels:
641 **AQ:** water, **B:** biofilm, **A:** August, **N:** November, **F:** February

642 **Figure 6:** Heatmap showing the taxonomic distribution of samples used to study biofilm
643 succession over time with relative abundance > 0.5%. **Others** indicate the abundance of
644 bacterial groups with < 0.5% of relative abundance.

645 **Figure 7:** **A)** Venn diagram showing shared and unique OTUs for surface and ground water
646 samples over time **B)** Shared OTUs between both types of water habitats SW and GW **C)**
647 metric multi-dimensional scaling (MDS) diagram **D)** table showing the results from the
648 analysis of similarity statistics (ANOSIM). Labels: **S:** surface, **G:** ground, **Ag:** August, **Nov:**
649 November, **Feb:** February, **3:** 3 months, **6:** 6 months, **12:** 12 months.

650 **Figure 8:** Alpha diversity metrics for biofilm samples used to study biofilm succession. **A)**
651 Shannon diversity index. **B)** Chao richness indicator. **C)** Dominance-H.

652

653 **Supplementary material:** Table showing Spearman's rank correlations with significant
654 statistical values

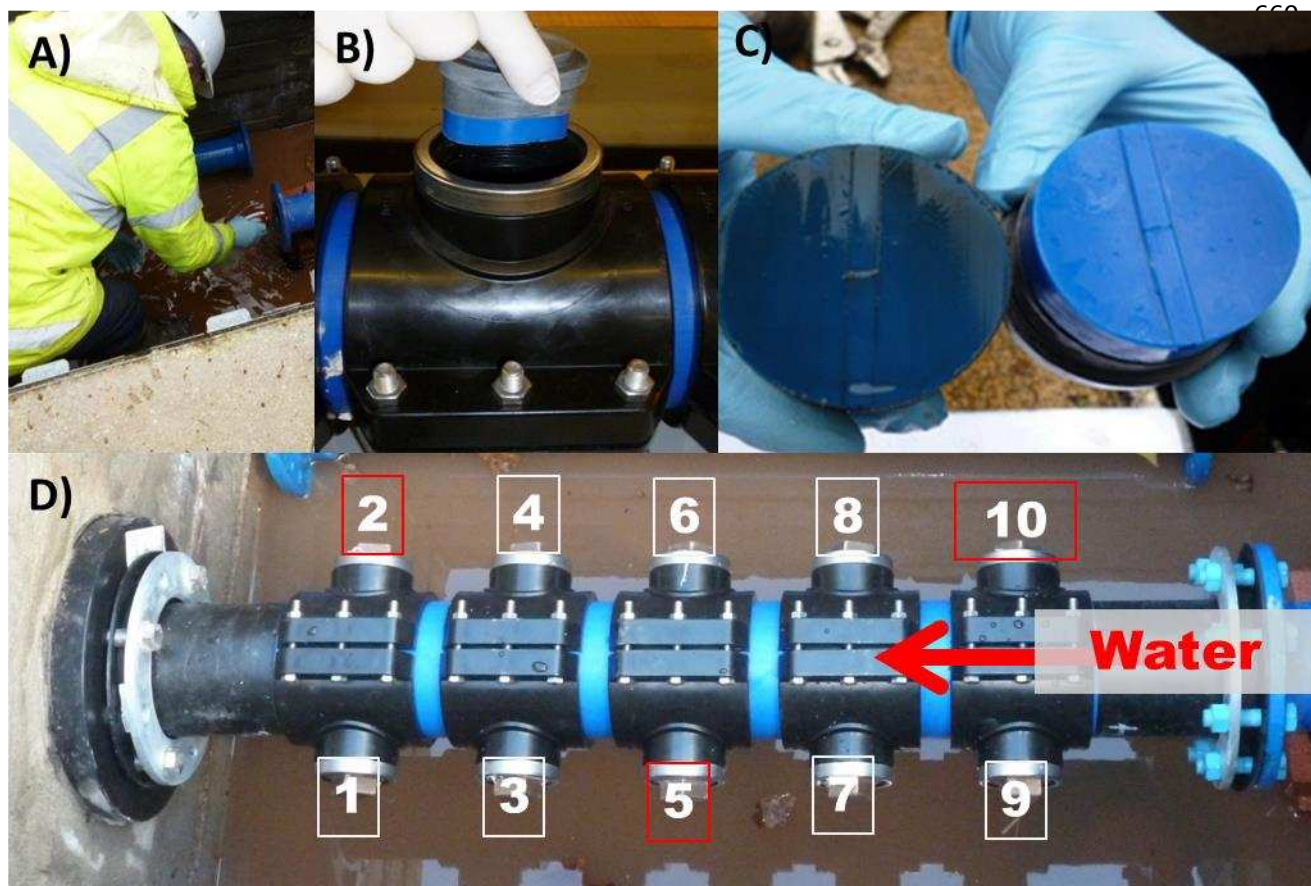
655

656 Table 1

	Surface Water			Ground Water		
	Aug-13	Nov-13	Feb-14	Aug-13	Nov-13	Feb-14
HPC 37°C (cfu/ml)	0	0	0	0	0	658
HPC 22°C (cfu/ml)	0	0	0	0	1.5±0.3	0
E coli (cfu/100ml)	0	0	0	0	0	0
Temperature (°C)	21.8±0	7.3±0.1	5.3±0	14.1±0	7.45±0.3	8.05±0
Turbidity (NTU)	0.12±0.01	0.05±0.1	0.2±0.03	0.12±0	0.1±0	0.1±0
Conductivity 20°C (uS/cm)	311±0	274.5±0.7	244.5±1.5	470±0	461.5±0.3	451.5±11.5
Conductivity 25°C (uS/cm)	344±0	303.5±0.7	270.5±1.50	520±0	510.5±0.3	499.5±12.5
pH	8±0	7.4±0	7.95±0.05	7.7±0	7.5±0.3	7.6±0
Alkal MO (mg CaCO3/L)	64.5±4.9	50±2.8	52.5±0.5	196.5±2.12	183.5±0.3	*
Ammonia as N (mg N/L)	0.02±0	0.07±0.1	<0.02	0.02±0	<0.02	<0.02
Tot oxid N (mg N/L)	0.7±0	1.2±0	1.4±0	8.55±0	8.3±0.3	9±0
Nitrite as N (mg N/L)	0.003±0	0.003±0	0.003±0	0.003±0	0.003±0	<0.003
Nitrate as N (mg N/L)	0.67±0	1.175±0	1.41±0.01	8.54±0.1	8.28±0	9.02
Ammonia (mg NH3/L)	0.01±0	0.07±0.1	<0.01	<0.01	<0.01	<0.01
Nitrite (mg NO2/L)	0.01±0	0.003±0	0.01	<0.01	<0.01	<0.01
Nitrate (mg NO3/L)	2.99±0.01	1.75±0	6.24	37.8±0.42	36.65±0.03	40±0
Orthophosphate (mg P/L)	<0.003	<0.003	<0.03	0.03±0	<0.03	<0.03
Sulphate (mg SO4/L)	76±0	61±0	47.5±0.5	24.5±0	24±0	16±0
Chloride (mg Cl/L)	21±0	20±0	19±0	18±0	19±0	17±0
Free Cl ₂ (mg Cl/L)	0.35±0	0.21±0	0.33±0	0.32±0	0.36±0	0.22±0
Total Organic Carbon (mgC/L)	0.95±0.07	0.01±0	0.95±0.05	0.95±0	0.6±0	0.65±0.05
Al mg/L	0.01±0	0.02±0	0.01±0	<0.01	<0.01	<0.01
Mn mg/L	0.002±0	0.002±0	0.0015±0	<0.001	<0.001	<0.001
Fe mg/L	0.01±0	0.02±00	0.02±0	<0.01	<0.01	<0.01
Cu mg/L	0.01±0	<0.01	<0.01	<0.01	<0.01	<0.01
Zn mg/L	0.01±0	<0.01	<0.01	<0.01	<0.01	<0.01

* No measured

659 Figure 1

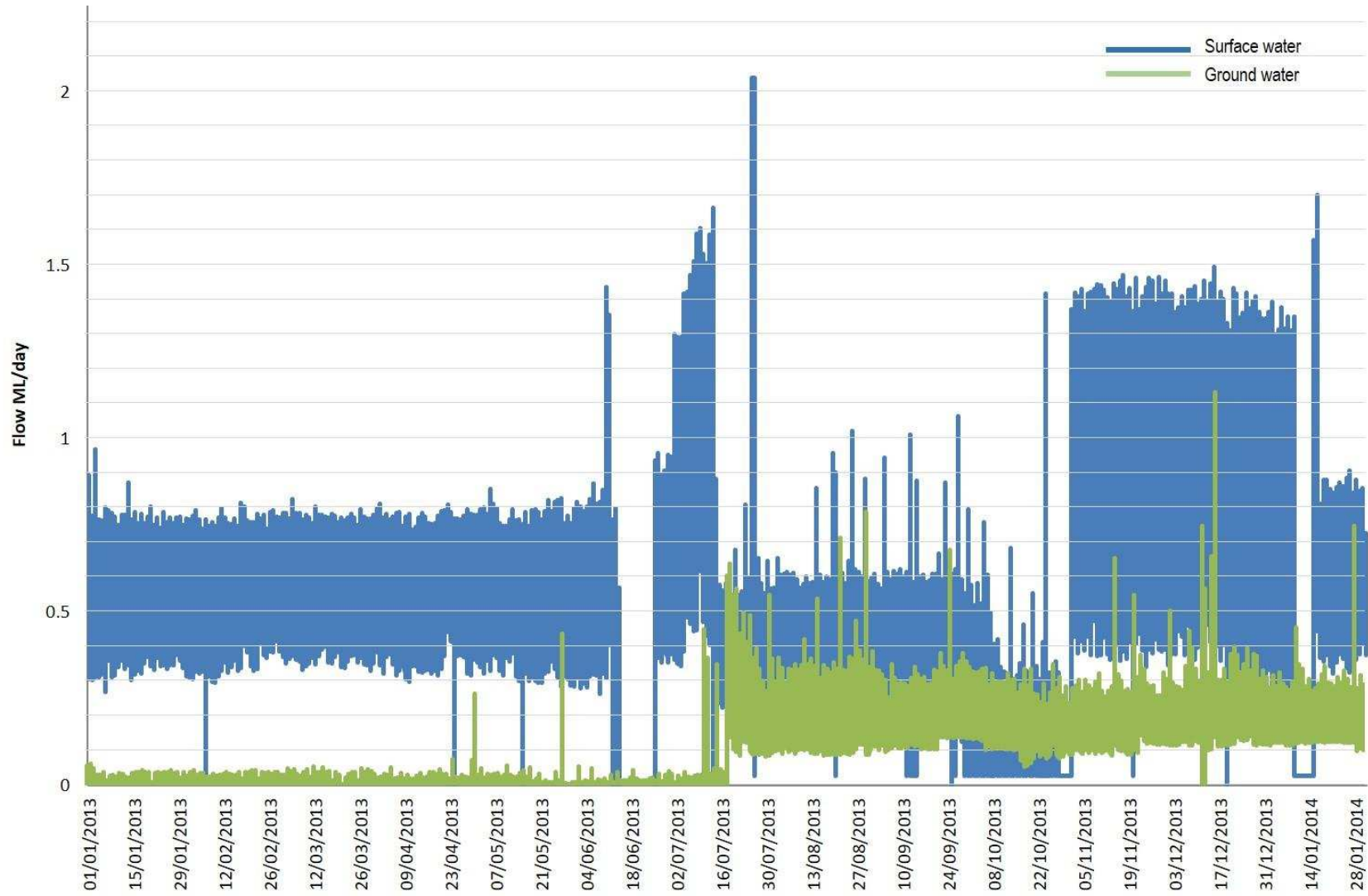


672

673

674

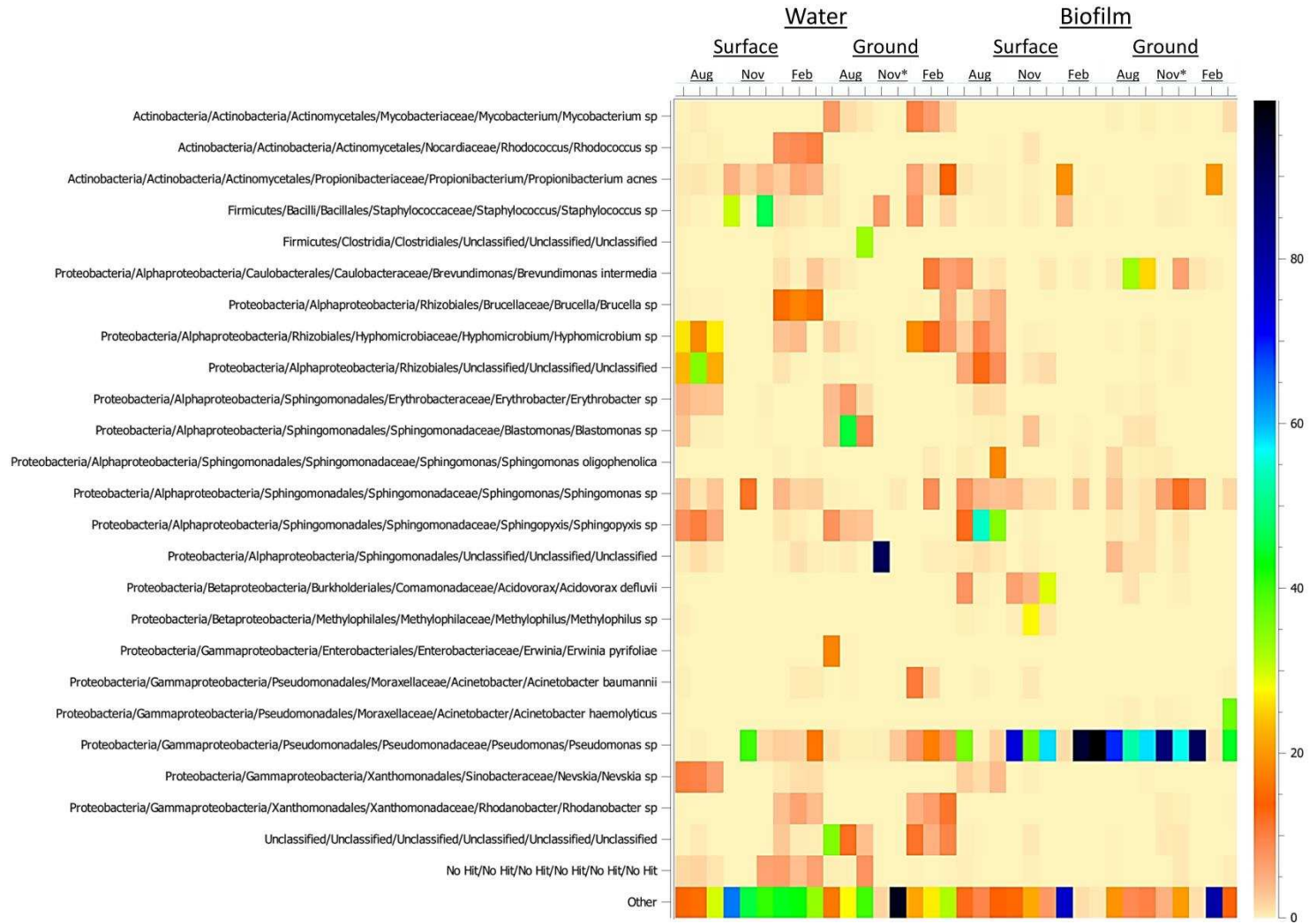
675 Figure 2



676

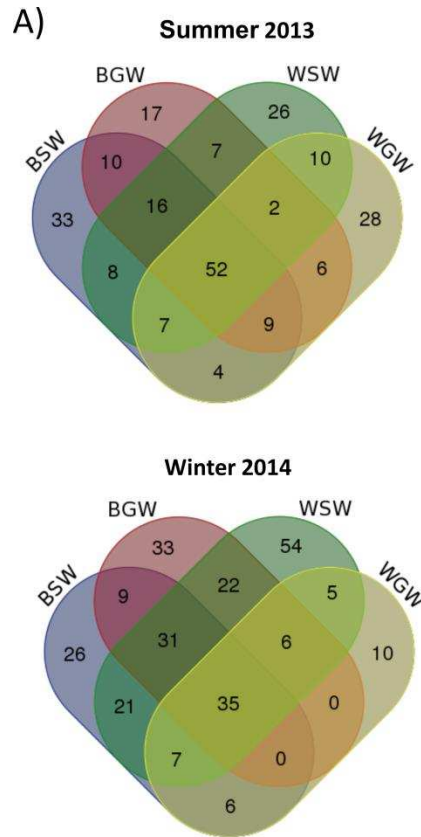
677

678 Figure 3

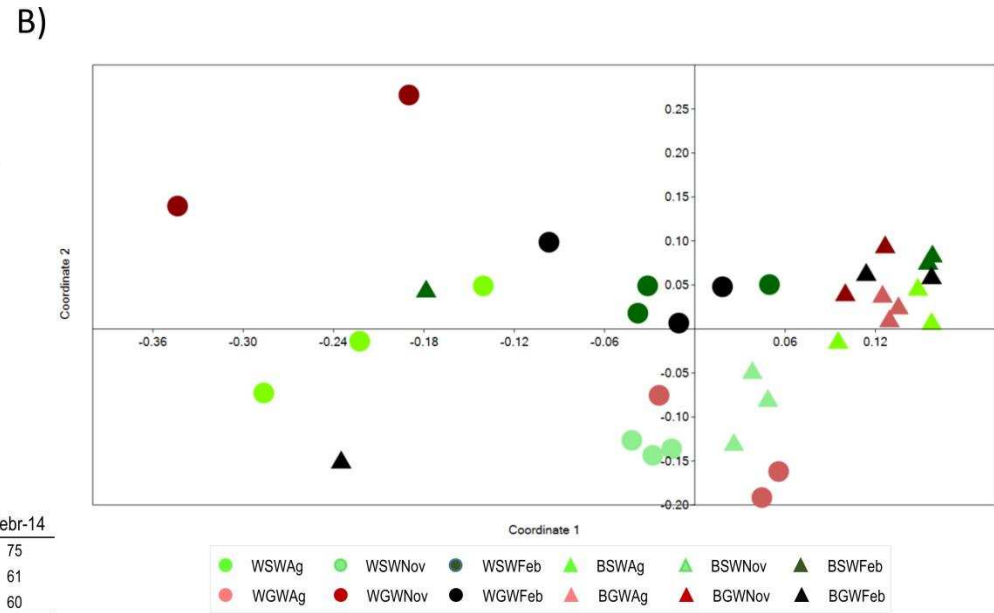


679

680 Figure 4



	Aug-13	Nov-13	Febr-14
Shared	87	11	75
BSW vs. BGW			
Unique to BSW	52	13	61
Unique to BGW	32	29	60
BSW vs. WSW			
Shared	83	7	94
Unique to BSW	56	33	41
Unique to WSW	45	18	87
BGW vs. WGW			
Shared	69	7	41
Unique to BGW	50	17	95
Unique to WGW	49	25	28
WSW vs. WGW			
Shared	71	11	53
Unique to WSW	57	14	128
Unique to WGW	47	21	16



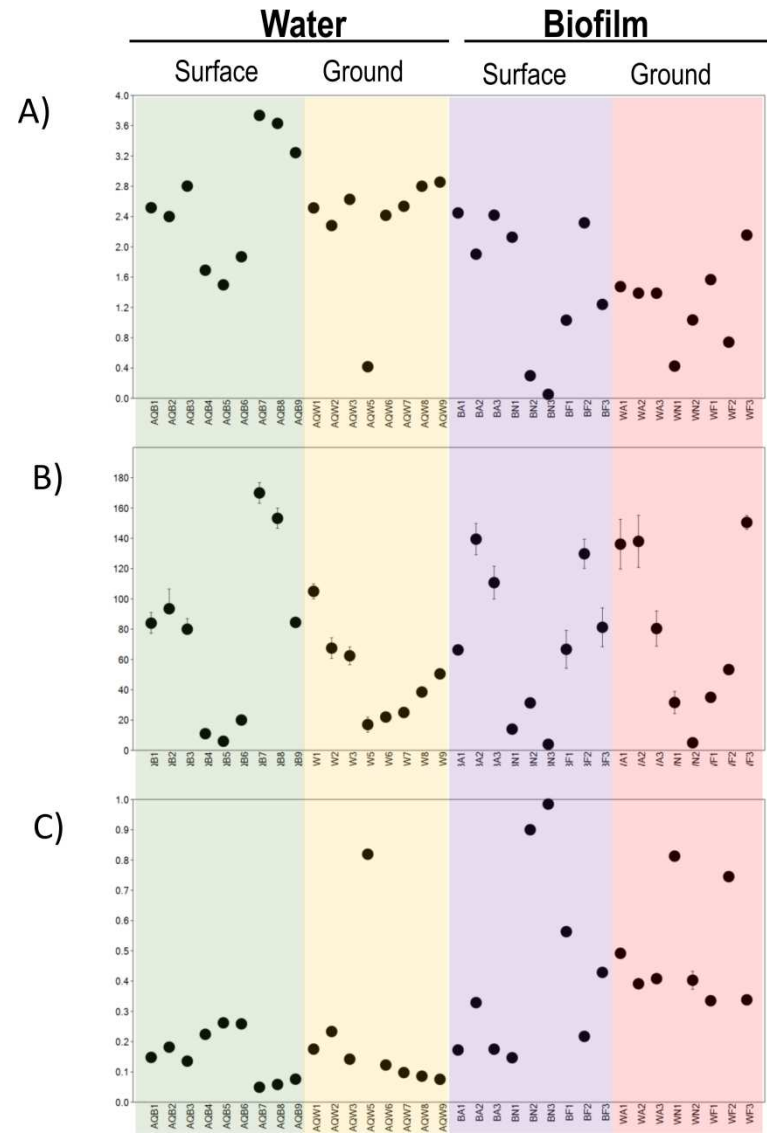
C)

	R values		
	WGW	BSW	BGW
WSW	0.1361*	0.2147*	0.4536**
WGW		0.2765**	0.3775**
BSW	0.2765**		0.04036
BGW	0.3775**	0.04036	

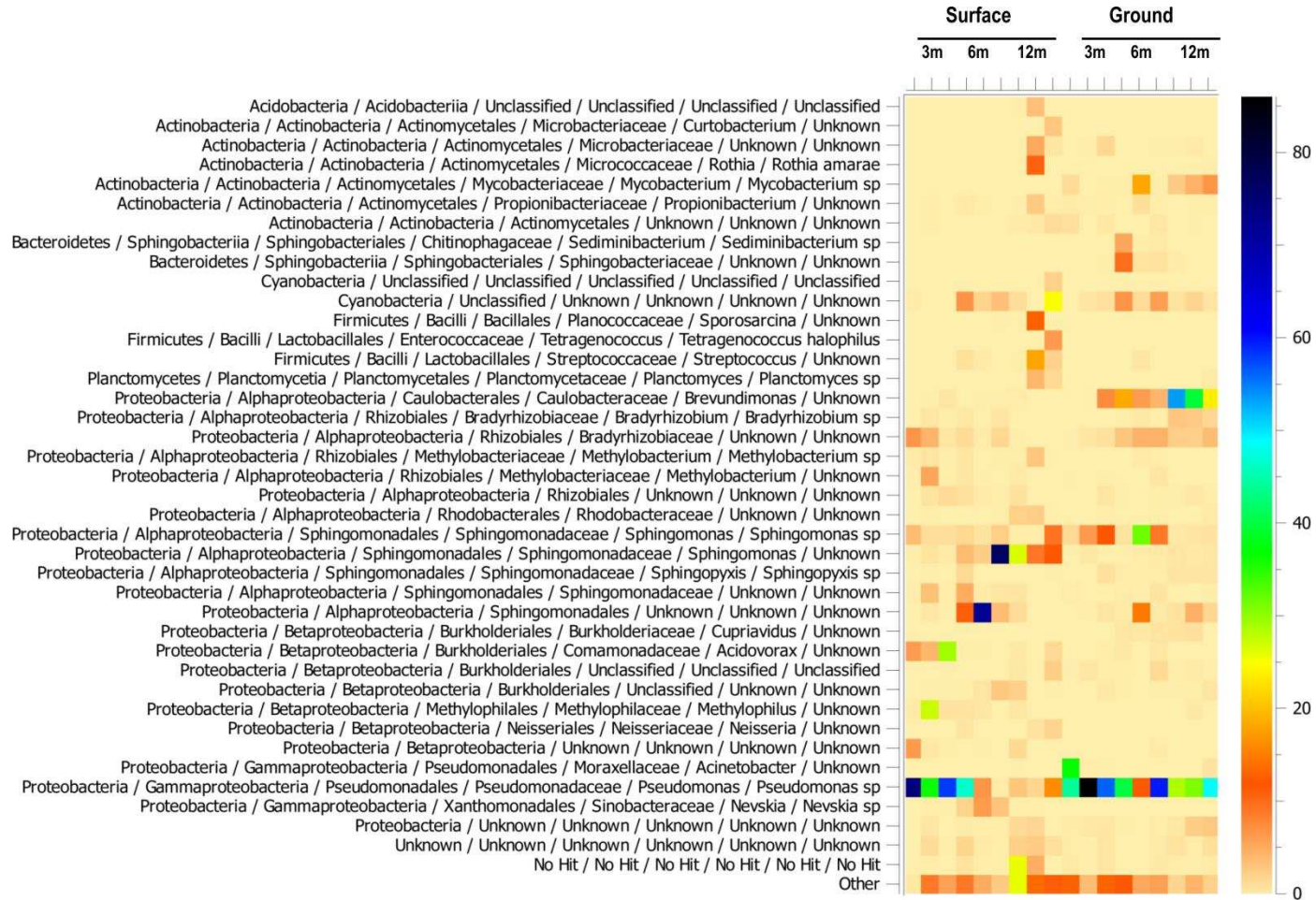
* p < 0.05 ; ** p < 0.01

681

682



685 Figure 6

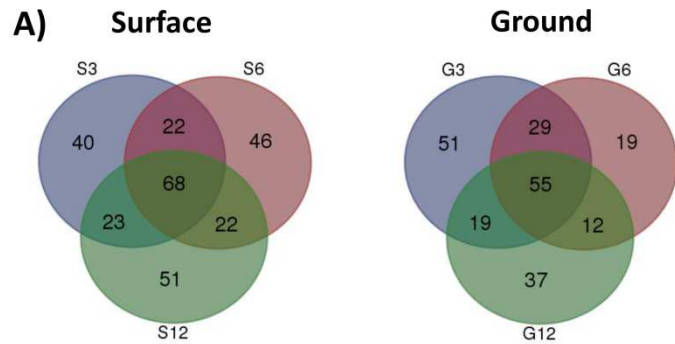


686

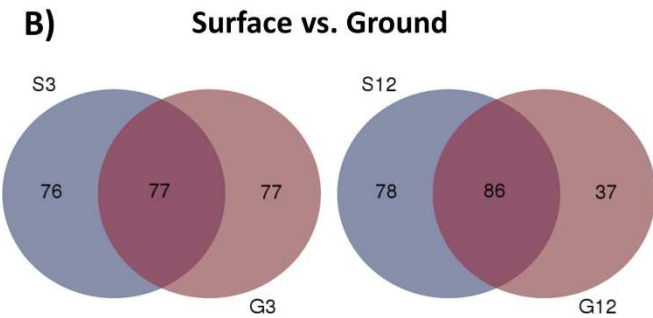
687

688

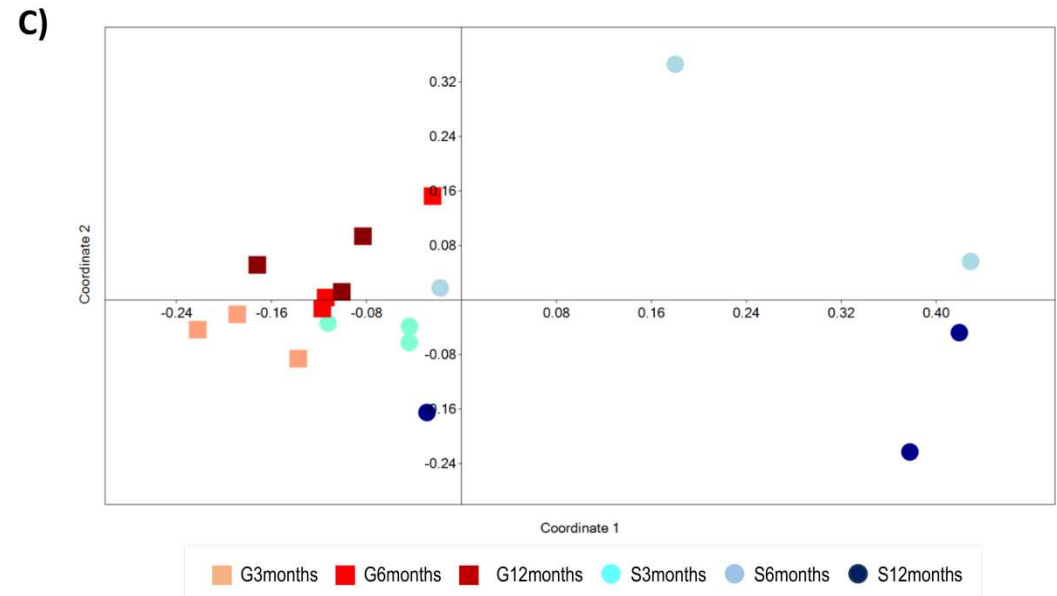
689 Figure 7



Surface			Ground		
S3 vs. S6	Shared	90	G3 vs. G6	Shared	70
	Unique to S3	63		Unique to G3	31
	Unique to S6	68		Unique to G6	84
S3 vs. S12	Shared	91	G3 vs. G12	Shared	80
	Unique to S3	62		Unique to G3	49
	Unique to S12	73		Unique to G12	74
S6 vs. S12	Shared	90	G6 vs. G12	Shared	48
	Unique to S6	68		Unique to G6	56
	Unique to S12	74		Unique to G12	67



Surface vs. Ground	
S3 vs G3	Shared: 77
	Unique to S3: 76
	Unique to G3: 77
S12 vs G12	Shared: 87
	Unique to S12: 78
	Unique to G12: 37



D) R Values

	WSWAg	WSWNov	WSWFeb	WGWAg	WGWNov	WGWFeb	BSWAg	BSWNov	BSWFeb	BGWAg	BGWNov
WSWAg	0.8889										
WSWNov	1	0.8889									
WSWFeb	0.8889	0.9259	1								
WGWAg	0.75	0.4167	0.5833	0.9167							
WGWNov	1	0.6296	0.4074	0.5926	0.75						
WGWFeb	0.963	0.963	1	1	1	0.8148					
BSWAg	1	1	1	1	1	0.6667	1				
BSWNov	0.4815	0.2963	0.3333	0.4444	0.5	0.1481	0.3333	0.2593			
BSWFeb	1	1	1	1	0.9167	0.6296	1	0.4444	0.4815		
BGWAg	1	1	1	1	1	0.6667	1	0.8333	-0.25	1	
BGWNov	0.3704	0.4444	0.3704	0.4815	1	0.1852	0.4074	0.1111	-0.0741	0.2963	0
BGWFeb											

** p < 0.01; * P < 0.05

690

691

