1 **16S rRNA assessment of the influence of shading on early-**
2 **successional biofilms in experimental streams**
3
4 Katja Lehmann[1]*, Andrew Singer[1]**, Michael J. Bowes[1], Nicola L. Ings[2], Dawn
5 Field[1], Thomas Bell[3]**
6
7
8 [1]NERC Centre for Ecology & Hydrology, Wallingford, OX10 8BB, UK
9 [2]Queen Mary University of London, London, UK
10 [3]Imperial College London, Department of Life Sciences, Silwood Park Campus, SL5
11 7PY, UK
12
13
14
15 *correspondence to
16 **contributed equally
17
18
19

20 **Abstract**
21

22 Elevated nutrient levels can lead to excessive biofilm growth, but reducing nutrient
23 pollution is often challenging. There is therefore interest in developing control
24 measures for biofilm growth in nutrient-rich rivers that could act as complement to
25 direct reductions in nutrient load. Shading of rivers is one option that can mitigate
26 blooms, but few studies have experimentally examined the differences in biofilm
27 communities grown under shaded and unshaded conditions. We investigated the
28 assembly and diversity of biofilm communities using *in situ* mesocosms within the
29 River Thames (UK). Biofilm composition was surveyed by 454 sequencing of 16S
30 amplicons (~400 bp length covering regions V6/V7). The results confirm the
31 importance of sunlight for biofilm community assembly; a resource that was utilized
32 by a relatively small number of dominant taxa, leading to significantly less diversity
33 than in shaded communities. These differences between unshaded and shaded were
34 either because of differences in resource utilization or loss of diatom-structures as
35 habitats for bacteria. We observed more co-occurrence patterns and network
36 interactions in the shaded communities. This lends further support to the proposal that
37 increased river shading can help mitigate the effects from macronutrient pollution in
38 rivers.

39

40 **Introduction**
41 Seasonal algal and cyanobacterial blooms have become regular occurrences in many
42 watersheds (Dodds, et al. 2002, Paerl, et al. 2011), and are predicted to increase in
43 frequency as a result of human population growth and climate change (Paerl, et al.
44 2011, Johnson *et al*., 2009). The increasing frequency of algal blooms in rivers
45 worldwide could have substantial economic and ecological consequences, there is
46 hence much interest in mitigating their impacts. The general consensus is that the
47 primary driver of algal blooms is concentration of macronutrients, which is increasing
48 due to growing human populations, agricultural intensification and increased
49 collection and release of urban wastewater (Bowes, et al. 2012, Mainstone and Parr
50 2002).
51 Options to reduce harmful blooms have focused on reducing effluent fluxes
52 (Kelly and Wilson 2004) (Neal, et al. 2010). The addition of tertiary treatment to

53 many sewage treatment works has led to substantial reductions in river macronutrient
54 contamination. However, reductions in macronutrient concentration are typically
55 costly, and do not always result in reductions in riverine biofilm (or periphyton)
56 growth or in improvements to other proxies of overall river health. What is more, the
57 relationship between harmful algal blooms and macronutrient enrichment is not
58 always a linear one. As Smayda (2008) notes, it is often difficult to trace harmful
59 algal blooms back to nutrient enrichment. In addition, some harmful blooms are
60 caused by nitrogen-fixers that thrive under nutrient-limited conditions (Paerl, et al.
61 2011). In rivers, catchment area, residence time and temperature are also important
62 factors (Desortova and Puncochar 2011, Bowes, et al. 2012). Notable examples where
63 nutrient management alone appears to be insufficient to control blooms are the Rivers
64 Thames in the UK and Berounka in the Czech Republic. In both rivers, annual means
65 of soluble reactive phosphorus concentrations have declined over the last decades
66 (from ca 1000 $\mu g\ l^{-1}$ to ca 200 $\mu g\ l^{-1}$ and ca 430 $\mu g\ l^{-1}$ to ca 160 $\mu g\ l^{-1}$ respectively).
67 These concentrations are still high enough for the algal biomass (measured as
68 chlorophyll-*a*) to remain similar to that observed before phosphorus mitigation
69 measures were introduced (Bowes, et al. 2012, Desortova and Puncochar 2011).
70 There is therefore a need for complementary controls that enhance the impact of
71 nutrient reductions.

72 One possibility is to use artificial or natural shading to reduce algae growth
73 rates in catchments with elevated nutrient concentrations. The rationale behind that is
74 that light can be as important in limiting growth of primary producers as
75 macronutrients (Rosemond 1993). Light-limited conditions can also prevent
76 dominance by a few fast growing species, particularly constrain the growth of low-
77 diversity communities consisting of filamentous species that can rapidly take
78 advantage of the high-nutrient environments and create a thick biomass mat that in
79 itself limits the growth of a range of organisms that grow in deeper biofilm layers
80 (Steinman 1992). In support of this, several experiments have shown that light
81 limitation mitigates the impact of eutrophication, even in nutrient rich environments
82 (Burrell, et al. 2013, Colijn and Cadée 2003, Hill and Knight 1988, Hill, et al. 1995,
83 Hutchins, et al. 2010, Mosisch, et al. 2001, Sanches, et al. 2011, Triska, et al. 1983,
84 Winterbourn 1990). Numerous studies have proposed a range of growth limiting
85 factors for controlling eutrophication. Most often cited is phosphorus, however, just

86 as important is flow velocity, grazing pressure, nitrogen pollution and light (McCall, 87 et al. 2014). "Natural" experiments have been particularly helpful in elucidating the 88 role of sunlight, whereby increased and decreased periods of natural sunlight on rivers 89 has directly translated into a corresponding increases and decreases in the intensity of 90 the resulting algal bloom (Read et al. 2014).

91 While there have been many studies on the effects of shading on overall 92 measures of biofilm growth, there has been relatively little research on how algal and 93 bacterial biofilm composition is affected by shading. The bacterial biofilm component 94 is less directly affected by light availability, but light levels could affect heterotrophs 95 through changes in UV radiation (Kahn & Wetzel 1999, Yoshikuni 2005, Thomas*, et* 96 *al*., 2009), or through indirect effects caused by changes to the autotrophic component 97 (Rier & Stevenson, 2002). This study uses an experimental approach to compare 98 riverine biofilm communities grown under shaded and unshaded conditions in the 99 River Thames. The Thames catchment is already heavily impacted by anthropogenic 100 activities. High and rising population density in the catchment are projected to put 101 additional pressure on water quality in the Thames, which might be exacerbated by 102 declining river flows and higher water temperatures brought about by climate change 103 (EA 2009, Evans, et al. 2003, Johnson, et al. 2009, Neal and Jarvie 2005). Algal and 104 bacterial biofilm communities were characterized using a molecular approach. Having 105 observed in previous experiments that increased algal growth prompted by excess 106 nutrients lead to biofilms which were dominated by few organisms and were less 107 diverse than those grown under nutrient-limiting conditions, we hypothesized that a 108 similar effect could be observed when light, another strongly limiting factor, was 109 restricted - as high light levels would allow a few fast-growing species from rapidly 110 dominating the communities. We tested two linked hypotheses:

111 1) biofilm grown in nutrient rich, shaded conditions assemble significantly different 112 biofilm communities than unshaded communities. Assuming that many members of a 113 biofilm community interact with each other in competitive, predatory or symbiotic 114 ways, any shift in one component of the population would lead to a shift in the others. 115 In that context, we assumed that 2) the biofilm communities that assemble in the 116 shade are more diverse than those that assemble in unshaded conditions, similar to 117 effects observed in nutrient experiments. The reason for the greater diversity might be

118 due to a reduced abundance of algal keystone species, which lead to a more complex
119 habitat (Bruno et al., 2003).
120
121
122 **Methods**
123 **Study site**
124 The study was conducted in experimental flumes placed within Seacourt Stream, a
125 side-branch of the Thames at Wytham in Oxfordshire, southern England
126 (Supplementary Figure 1; 51.786413, -1.317073 Lat/Long, Decimal Geographic
127 Coordinates). Seacourt Stream is a disused millstream directly fed by the Thames
128 (100m upstream). The site was selected due to its lack of natural shading.
129 Macronutrient composition at the start of the experiment was similar to that in the
130 main Thames branch, and showed little change during the study period (7th to 17th
131 September 2010; see Results).
132
133 **Experiment methodology**
134 We installed 12 experimental flow-through flumes at the study site (Supplementary
135 Figure 2). The mesocosms are described in Bowes *et al*. (2010: 384-9). Briefly, they
136 are constructed as blocks of 3 flumes that float directly in the river, allowing river
137 water to flow through freely. Each flume has a gate to standardise flow rates at the
138 upstream end and a sump to collect river debris ¾ of the length from the inlet (Bowes,
139 et al. 2010). The flumes are made of polyvinyl chloride sheeting set in an aluminium
140 frame, with each flume measuring 5 m x 0.3 m. For this experiment, 4 sets of flumes
141 (i.e. 12 flumes in total) were tethered to the riverbank and positioned 0.5 m above the
142 streambed, held by scaffolding poles. Floats ensured a constant water depth of ~6 cm
143 inside the flumes. The gap between the riverbed and mesocosms limits invertebrate
144 colonizers from entering the flumes.
145 Before the start of the experiment, we measured midday light intensity with a
146 SunScan SS1 light probe in both direct sunlight and full tree shade. In full tree shade,
147 the intensity of direct sunlight was reduced by 71%. We used layers of greenhouse
148 shading mesh, positioned directly above the flumes, to create light intensities
149 equivalent to those measured in full tree shade over parts of the flumes. Dividing each
150 channel in half, we grew shaded and unshaded biofilm next to each. A similar

151 approach has been used in other studies investigating the effect of light on biofilm
152 (Hill, et al. 2009, Hill, et al. 2011, Rier, et al. 2006). We positioned temperature
153 loggers (iButton, Maxim, San Jose, CA, USA) next to each set of tiles to determine
154 whether the shading also reduced water temperature. We set the flow rates in each
155 channel to 0.1 m s^{-1} (measured using a Valeport 801 flow meter) at the start of the
156 experiment.
157
158 **Sample collection and DNA sequencing**
159 We grew biofilm on 2 x 2 cm limestone tiles anchored to the bottom of the channels
160 for 9 days. This might have lead to a community that was not fully mature, but when
161 left longer, biofilms in previous experiment invariable sloughed off the tiles and
162 floated downstream. In the upstream half of each flume, there were 3 shaded tiles and
163 3 unshaded tiles. On day 9, we harvested the biofilms in all flumes and extracted the
164 DNA, pooling the 3 tiles within each treatment/flume. Briefly, we added 300 µl of
165 lysis buffer (100 mmM NaCl, 500 mM Tris (pH 8), 10% (w/v) sodium dodecyl
166 sulfate, 2 mg ml^{-1} proteinase K, 2 mg ml^{-1} lysing enzyme mix (both Sigma-Genosys,
167 Gillingham, UK)) and 300 µl of NaH_2PO_4 (pH 8.0) to the pooled sample, incubated
168 the DNA in a 55°C water bath for 30 min and mixed every 10 min., added 80 µl of
169 prewarmed 10% CTAB solution (65°C), incubated in 65°C for 10 minutes, added
170 680µl chloroform:isoamyl alcohol (24:1 vol/vol). The tubes were centrifuged for 5
171 minutes at 14000 rpm. The aqueous top layer was aspirated into a new tube and the
172 DNA precipitated by adding 300% (w/v) TE Buffer, pH 8.0 (10 mM TRIS-HCl, 1
173 mM EDTA, pH 8.0) and 200% (w/v) PEG/$MgCl_2$ mix (30% (w/v) PEG 8000, 30 mM
174 $MgCl_2$), leaving the samples overnight at 5°C (Paithankar and Prasad 1991). We then
175 centrifuged the replicates (12 per treatment) for 10 min at 14000 rpm, discarded the
176 supernatant and washed the DNA pellets by adding 300 µl 70% chilled ethanol. We
177 centrifuged the tubes again, discarded the ethanol and left the tubes to dry in a laminar
178 flow cabinet until the ethanol had evaporated. We added 50 µl ultrapure water and left
179 the DNA to resuspend for 1 h on the bench. We used the 454 GS-FLX TITANIUM
180 platform (Roche 454 Life Sciences, Branford, CT, USA) to produce tag-encoded 16S
181 amplicons of ~400 bp length. We targeted a fragment of the 16S ribosomal RNA
182 (rRNA), comprising the V6 and V7 regions. Primers used for the PCR were 967F, 5'-
183 CNACGCGAAGAACCTTANC-3', and 1391R, 5'- GACGGGCGGTGTGTRCA-3'

184 (Huber, et al. 2009: 1292-302, Huse, et al. 2008). These universal primers are
185 designed to amplify a large variety of 16S sequences, but, as with all universal
186 primers, it cannot be excluded that some OTUs (both chloroplast and bacterial 16S
187 rRNA) did not get captured. The sequencing libraries were generated through a one-
188 step PCR with a total of 30 cycles, a mixture of Hot Start and Hot Start high fidelity
189 taq polymerases and amplicons extending from the forward primers. DNA
190 amplification and pyrosequencing were carried out at Research and Testing
191 Laboratory (Lubbock, TX, USA).

192 **Bioinformatics**

193 We used CloVR 1.0 RC4 (Angiuoli, et al. 2011) on the Data Intensive Academic Grid
194 (DIAG, University of Maryland, USA) to run the QIIME workflow
195 'pick_otus_through_otu_tables.py' (Caporaso, et al. 2010). Within the QIIME
196 workflow: (1) we set the minimum and maximum sequence length to 100 bp and
197 2000 bp respectively, the maximum homopolymer length to 8 bp and maximum
198 number of ambiguous calls to zero; (2) clustering was performed using UCLUST with
199 a nucleotide sequence identity threshold within each cluster at 97% and alignment
200 against the Greengenes 16S database with PyNAST; (3) taxonomy assignment of each
201 OTU-representing sequence through the RDP classifier with a confidence threshold of
202 0.8. After quality control, the data set consisted of 101,617 combined reads for all 12
203 flume channels of the experiment. Clustering and chimera removal left 97,065
204 combined reads. Following from earlier studies (Lindemann, et al. 2013, Pillet, et al.
205 2011) we used the chloroplast 16S rRNA to focus on the algal communities.
206 Therefore, of those OTU's that were identified to Genus level, we divided the
207 community into algal-derived chloroplast reads and bacterial (including
208 cyanobacterial) reads. We equilibrated the number of sequences per sample by
209 randomly sampling without replacement (Hamady and Knight 2009, Koren, et al.
210 2013), resulting in 290 algal sequences per sample, and 732 bacterial sequences per
211 sample. OTUs that are discussed on the species level were blasted individually against
212 the RDP database (Cole, et al. 2009). Only fragments that could be matched at 97% or
213 above were classified to species level. The rarefied OTU tables were imported into
214 Primer (PRIMER-E Ltd, Ivybridge, UK), MEGAN (Huson, et al. 2007) and R for
215 further analysis.

216

217 **Statistics**

218 We were interested in how shading altered biofilm community composition and
219 diversity. We calculated diversity, tested that the diversity data was normally
220 distributed, then compared diversity across the treatments using analysis of variance
221 (ANOVA). We included both treatment and channel as factors. We added the location
222 of the flumes within the river channel as an additional factor/error, because not all
223 flumes could be placed next to each in the river channel and location effects could not
224 be excluded. We compared dissimilarity in community composition by calculating
225 Bray-Curtis dissimilarities between pairs of communities (Bray and Curtis 1957). We
226 tested for differences between treatments using PERMANOVA, a multivariate
227 permutation test analogous to ANOVA (Anderson 2005). The PERMANOVA design
228 was two-factorial, including treatment and channel. We used non-metric
229 multidimensional scaling (NMDS) (Kruskal 1964) to visualize differences between
230 the communities. We then used similarity percentages (SIMPER) to explore the
231 contribution of each species (Clarke 1993). SIMPER assesses the contribution of each
232 species to the observed dissimilarity between communities. PCR-based data cannot be
233 used to accurately assess abundances in the original samples, but given that all
234 samples were amplified in the same way, we assume that abundances can be
235 compared between our samples (but not with samples from other datasets). Lastly, we
236 used network analysis of co-occurrence patterns (Barberan, et al. 2012) to explore
237 possible connections between biofilm components. We used network analysis
238 implemented in the MEGAN software package (Huson, et al. 2007) to visualise co-
239 occurrence patterns in our data. The visualisation connects OTUs (here at the
240 taxonomic level of Class) that exceed a prescribed probability of co-occurrence. We
241 set the following threshold values: the threshold required for a taxon to be considered
242 present in a sample was 0.5%; the minimum and maximum percentage of samples in
243 which a taxon could occur was set to 15% (2 samples) and 100% (12 samples); the
244 minimum probability that a co-occurrence between two taxa had been observed was
245 set to 95%. All of the statistical results and figures were produced using CloVR
246 (Angiuoli, et al. 2011), MEGAN or the R environment (www.r-project.org).
247 Significance thresholds of $p < 0.05$ were used throughout to validate the results.

248

249 **Results**

250 The taxonomic distribution of the data shows that a large proportion of taxa were of
251 algal origin (Fig. 3). 19% of total number of sequences could not be identified to
252 Genus level, but only 1% of the 44% that were identified as algae were not
253 identifiable to Genus level. The Shannon diversity (Fig. 4) of both the algal and
254 bacterial communities was significantly higher in the shaded communities ($F_{1,1}$ = 36.4
255 (Algae) and 7.1 (Bacteria), p = 1.26e-04 (Algae) and 0.02 (Bacteria)). In the algal
256 component of the biofilm, the community was dominated by *Amphora sp. C10*,
257 *Melosira varians* and *Amphiprora paludosa str. CCMP 125 C52*, which accounted for
258 62% of overall relative abundance (Fig. 5). In contrast, in the shaded community
259 these 3 species were still dominant but accounted for only 44% of the overall relative
260 abundance (Fig. 5). For the bacterial component of the unshaded communities, the
261 most abundant species were *Curvibacter sp. str. HMD2015* (2.4%) and
262 *Steroidobacter sp. str. ZUMI 37* (2.4%). These two species were also the most
263 abundant in the shaded community (3.4% and 1.7%), but in both cases they accounted
264 for only approximately 5% of the total community (Fig. 5). Applying Pielou's
265 evenness measure to the samples confirms that the bacterial components of the
266 biofilm were significantly more even in composition than the algal ones, in both the
267 unshaded and shaded replicates (0.95 for the bacterial components, 0.7 for the algal
268 component, F=460, p< 2e-16).
269 Ordination of the communities using NMDS (Fig. 6) indicates distinct clusters of
270 unshaded and shaded communities for both the eukaryotes (PERMANOVA:
271 eukaryotes, Pseudo F = 7.60, p = 0.002 and bacteria, Pseudo F = 2.52, p = 0.006).
272 The unshaded communities are less variable than the shaded ones.
273 We used SIMPER analysis (Clarke 1993) to investigate which OTUs
274 contributed most to the observed dissimilarity between the shaded and unshaded
275 assemblages. In the algal component, the three most abundant diatoms also
276 contributed most to the observed dissimilarity: *Amphora sp. C10* contributed 8.6%,
277 *Melosira varians* 8.4% and *Amphiprora paludosa str. CCMP 125 C52* 6.5%. All three
278 were more abundant in the unshaded treatment and together accounted for 23% of the
279 overall observed difference. The next six important algal contributors, however, were
280 more abundant in the shaded treatment than in the unshaded treatment, and accounted
281 for 14% of the overall observed difference. Due to the greater evenness of the
282 bacterial biofilm component, the 10 major contributors in the bacterial replicates only

283 accounted for 6% of the overall dissimilarity, with the two most abundant species
284 (*Steroidobacter sp. str. ZUMI 37* and *Curvibacter sp. str. HMD2015*) contributing just
285 2% of dissimilarity.
286
287 We used co-occurrence network analysis to explore the relationship between bacterial
288 and algal taxa (Fig. 7). The figure shows one main network in the unshaded samples,
289 consisting of common bacteria and diatoms found in all twelve replicates (Fig. 7:
290 *Bacillariophyceae, Gemmatimonadetes, Rhodobacter, Bacteroidetes,*
291 *Prosthecobacter, Acidobacteria, Anaerolineae*). This major network was also
292 detectable in the shaded replicates, consisting of four of the nodes that were present in
293 the unshaded samples (Fig. 7: *Bacillariophyceae, Bacteroidetes, Acidobacteria,*
294 *Anaerolineae*). These were joined by four additional nodes, differing from the ones in
295 the unshaded samples (Fig 7: *Methylobacter, Variovorax, Polaromonas,*
296 *Planctomycetes*). That means this network included one additional member in all
297 twelve replicates of the shaded treatment.
298 In eleven of the unshaded replicates there was also a small network consisting
299 of just two nodes, *Frateuria* and *Planctomycetes*. This 2-node-network containing
300 *Frateuria* was also detectable in ten of the shaded samples, but the *Planctomycetes*
301 node was replaced by *Verrucomicrobia*.
302 In eleven of the shaded samples, there were also two small networks that were
303 absent in the unshaded samples (*Actinobacteria* and *Nitrospira, Rhodobacter* and
304 *Rhodoferax*), and in two of the replicates there was a network consisting of
305 *Acidovorax* and *Shewanella*.
306 Throughout the experiment, Seacourt Stream had nutrient concentrations of
307 around 234 μg l^{-1} soluble reactive phosphorus (SRP), 5.18 mg l^{-1} nitrate (N), and the
308 dissolved reactive silicon 3.05 mg l^{-1}, which is considered high for SRP, low for N,
309 and below average for silicon (EA 2012, Neal, et al. 2005). Silicon is typically
310 depleted at periods when diatoms are ‘in bloom’, which is consistent with this period
311 of study. The water temperature of the shaded areas was at all times identical to that
312 of the unshaded areas and in all channels (averaging at 15.5°C) throughout the
313 experiment.
314
315 **Discussion**

316 This study joins a number of other studies in observing significant changes to algal
317 assemblages under reduced light conditions (Bowes, et al. 2012, Guariento, et al.
318 2011, O'Driscoll, et al. 2006). Uniquely, however, our results show that reduced light
319 conditions affect both algal and bacterial components of biofilm.

320 A marked result of shading is a change of dominance in the most prevalent
321 organisms. Whilst PCR can skew abundance-patterns found in the original sample,
322 the decrease in dominance of the three most dominant diatoms, *Amphora sp. C10,*
323 *Melosira varians,* and *Amphiprora paludosa str. CCMP 125 C52,* under shaded
324 conditions is consistent with findings by Hill, et al. (2011), who note in addition, that
325 light is a more limiting factor for autotrophs than nutrient availability. In another
326 study, Sanches, et al. (2011) confirm and expand on these findings by showing that
327 low light availability does not only limit biofilm growth, but also nutrient propagation
328 within the biofilm, thereby affecting the autotrophic to total biofilm biomass ratio. In
329 the Sanches *et al*. experiment, autotrophic biomass was highest under high light
330 conditions and N enrichment, whereas heterotrophic biomass increased under
331 enrichment for both P and N (Sanches, et al. 2011). To relate that back to the Seacourt
332 Stream experiment: the macronutrient content in Seacourt Stream measured at the
333 beginning of the experiment was sufficiently enriched enough to allow for an increase
334 in abundance of autotrophs. We have shown this increase in a parallel experiment,
335 which measured biomass and chlorophyll a content conducted in the downstream area
336 of the same flumes (published as Bowes *et al*., 2012b). The parallel experiment
337 certainly showed higher algal biomass accrual in the unshaded treatment, and that is
338 probably reflected in the observed increase of the dominant diatom. Bowes *et al*.
339 (2012) further support the findings of this study in as much as the biomass tiles of the
340 shaded treatment showed greater algal diversity than the unshaded tiles.

341 The changes in the bacterial community components were less marked than
342 for the algae, but again the most visible pattern was a reduction in the relative
343 abundance of the most prominent dominant organisms in the shaded samples. One
344 potential explanation is that this is driven partly by the correlated changes in the algal
345 community. A possible reason for this is that specific bacteria utilize organic
346 compounds excreted by algal species (Ylla, *et al*. 2009). A greater diversity of algae
347 might provide a higher diversity of exudates, which would in turn support a higher
348 diversity of heterotrophs. Kritzberg, et al. (2006) have stated that 30-65% of bacterial

349 production in lakes is supported by algae-derived autochthonous carbon. A reduction
350 in algae due to less light is also likely to have lead to a reduction in autochthonous
351 carbon availability on the tiles, to which the bacterial community responded with a
352 shift in structure. Chang (2010) has hypothesized that such structural changes could
353 be due to a shift from heterotrophic consumers, to consumers that are more likely
354 adapted to allochthonous carbon sources. Finally, many bacteria colonize the
355 exoskeletons of diatoms and their diffusive boundary layer, a thin layer of fluid
356 directly surrounding the diatom known as the 'phycosphere' (Bell and Mitchell 1972,
357 Rier and Stevenson 2002, Znachor, et al. 2012). This layer contains extracellular
358 products produced by the diatom, and bacteria living in the vicinity might provide the
359 diatom with products it cannot produce itself (Amin, et al. 2012). A reduction of
360 Baccilariophyceae numbers would also reduce the number of such colonizers. The co-
361 occurrence network analysis suggests, that there were indeed a number of bacterial
362 classes in our samples with occurrence patterns that matched those of the
363 Baccilariophyceae in the experiment.

364 It is not possible to infer much about the nature of taxon interactions from our
365 data, but the result of the co-occurrence analysis could suggest that the influence of
366 light is so great, that interactions between taxa are less important when enough light is
367 available. To expand on that, under shaded conditions, a reduced amount of freely
368 available DOM and other metabolites might make it more important for biofilm
369 organisms to interact with co-habitants of the biofilm, as possibly indicated by the
370 increase of co-occurrence networks in the shaded replicates. At the same time, light
371 limitation might forge relationships that differ from those in light non-limited
372 conditions, as seen in the changing *Frateuria*-networks. A possible cause for the
373 network formations observed between *Frateuria* and other organisms could be that
374 *Frateuria* is unable to synthesize some of the compounds required for its growth
375 (Hashidoko 2005). In the unshaded samples, *Frateuria* appears to form a network
376 with *Planctomycetes*, but in the shaded samples, *Planctomycetes* are part of the main,
377 bigger network, whereas *Frateuria* forms a network with *Verrucomicrobia*.

378 The *Verrucomicrobia* are found in a greater number of the shaded samples than the
379 unshaded samples, but it is probably more relevant that in the shaded samples, the
380 *Planctomycetes* are part of the main network. One possible cause for this is that the
381 shaded and unshaded replicates harbor different species of *Planctomycetes* and

382 *Verrucomicrobia*. Another cause could, however, be that in the shaded samples, the
383 *Planctomycetes* rely on a close relationship with other organisms to obtain products,
384 which under non-limited light conditions, are easily available, leaving *Frateuria* to
385 seek out another organism to form a relationship with. Any statement on the nature of
386 the exchanges is purely speculative, but it may be interesting to note that *Frateuria*
387 have been linked to methanogens in the past (Romanovskaya and Titov 1992), and
388 that both *Verrucomibrobia* and *Planctomycetes* include methanogens amongst their
389 groups of species (Chistoserdova, et al. 2004, Dunfield, et al. 2007). A possible
390 outcome of such a change in interactions could be that there are functional differences
391 between shaded and unshaded communities.

392 All of the dominant diatom species in this study can cause blooms (Hillebrand
393 and Sommer 1997, Ohtsuka 2005, Khare and Chaurasia 2009, Vanelslander, et al.
394 2009, Dorigo, et al. 2010, Paerl, et al. 2011). Even though we observed that a shaded
395 environment led to less dominance of diatom taxa, it cannot be excluded that these
396 diatom taxa could adjust to shady conditions by reaching saturation levels at lower
397 light intensities, as previously described (Rier, et al. 2006). Notably, there were two
398 possible blooming species, *Haslea nipkowii* and *Synedra hyperborea str. CCMP* that
399 were more abundant under shaded than under unshaded conditions. Likewise, 15 of
400 the 29 observed cyanobacteria species were more abundant under shaded conditions,
401 too. This means that riparian shading can only be one tool in managing algal and
402 cyanobacterial blooms.

403 Lastly, it is important to consider what effect riparian shade has on other riverine
404 organisms. Invertebrates, for example, have been shown to decline when shading is
405 increased by more than 60% (Quinn, et al. 1997). Hence, mitigating measures for
406 'algal' blooms, such as shading, could have wider ecosystem implications.

407 There was a surprisingly low abundance of algal genera other than diatoms. It
408 is possible that the experiment was stopped before a significant number of
409 filamentous algae could establish themselves. Inferences drawn in this study might
410 therefore only be applicable to early succession biofilms. Alternatively, the low
411 number of filamentous species could have been the result of the exclusion of snails
412 from the flumes, which selected for diatoms (Rosemond 1993). It is also unclear
413 whether the universal primers that were used to amplify our 16S sequences might
414 have been more suitable for diatoms than for filamentous algae (Chung and Staub

415 2003). The shaded replicates had two outliers (S1, S2), which probably received more
416 sun during the experiment due to the sun's angle at particular times of the day,
417 however, these did not affect the statistical significance of the overall results.
418

419 **Conclusions**

420 Shading has a marked effect on the structure and diversity of both algal and bacterial
421 assemblages in biofilm. It significantly reduces the prevalence of diatoms that are
422 known to cause nuisance blooms under nutrient-enriched conditions, and creates
423 communities that are more even and diverse. Our algal results support findings e.g. by
424 Hill, et al. (2009), Ghermandi, et al. (2009) and Bowes, et al. (2012) that suggest
425 riparian shading is an effective tool in controlling biofilm growth rates and managing
426 the effects of eutrophication. Whilst it may not seem practicable to have extensive
427 riparian planting schemes, the need to mitigate climate change might make such
428 schemes more palatable. Recent management practice in the UK had begun to
429 advertise how to create riparian shade where it is absent (Lenane, 2012). The obvious
430 advantages that shading has in reducing algal blooms and keeping water temperatures
431 low (Warner and Hendrix, 1984, Lenane, 2012) seem to make the planting of shading
432 desirable even if it is a longer term project (Lenane, 2012).

433 Even if, however, riparian shading presents itself as a useful tool to manage eutrophic
434 streams that experience blooms, more research needs to be conducted to assess if
435 shading becomes ineffective as diatoms adjust to lower light levels (Rier, et al. 2006)
436 or are replaced by species better suited to shade. It is also necessary to investigate if
437 shading has a negative effect on the function of biofilm and on invertebrate grazers.

438 To understand the observed differences between the communities, it would be
439 useful to investigate any changes in function. This could be an examination to
440 determine if bacterial assemblages in shaded rivers are less equipped to process
441 glycosate, which is produced by periphytic algae, or simply transcriptomic analysis of
442 the whole communities, as transcriptomics would identify if the communities are
443 functionally different. It would also be interesting to test whether the diverse and even
444 communities created by riparian shading prove to be more resilient to stress and
445 resistant to pollution events. In that context, it should be tested what effect different
446 community assemblages have on nutrient cycling and biofilm function. Another
447 question is to what degree changes to biofilm nutrient stoichiometry (Cross, et al.

448 2005) cause changes to higher trophic levels. Whilst shading might shift biofilm
449 community structure in such a way that harmful blooms are reduced, it might produce
450 unexpected effects on higher trophic levels in the river.

451

452 **Acknowledgements**

453 Many thanks to Alan Grafen for his advice on the statistical methods. This work was
454 funded by the NERC Centre for Ecology & Hydrology National Capability funding
455 through the CEH Thames Initiative and the NERC–Knowledge Transfer (PREPARE)
456 Initiative contract NE/F009216/1.

457

458 **References**

459 Amin SA, Parker MS, Armbrust EV. Interactions between Diatoms and Bacteria. *Microbiology and Molecular*
460 *Biology Reviews* 2012;**76**: 667-84.
461 Anderson M. PERMANOVA: a FORTRAN computer program for permutational multivariate analysis of
462 variance. *Department of Statistics, University of Auckland, New Zealand* 2005;**24**.
463 Angiuoli S, Matalka M, Gussman G *et al*. CloVR: A virtual machine for automated and portable sequence analysis
464 from the desktop using cloud computing. *BMC Bioinformatics* 2011.
465 Barberan A, Bates ST, Casamayor EO *et al*. Using network analysis to explore co-occurrence patterns in soil
466 microbial communities. *ISME J* 2012;**6**: 343-51.
467 Bell W, Mitchell R. Chemotactic and Growth Responses of Marine Bacteria to Algal Extracellular Products. *Biol*
468 *Bull* 1972;**143**: 265
469 Bowes MJ, Gozzard E, Johnson AC *et al*. Spatial and temporal changes in chlorophyll-a concentrations in the
470 River Thames basin, UK: Are phosphorus concentrations beginning to limit phytoplankton biomass?
471 *Science of The Total Environment* 2012a;**426**: 45-55.
472 Bowes MJ, Ings NL, McCall SJ *et al*. Nutrient and light limitation of biofilm in the River Thames: Implications for
473 catchment management. *Science of The Total Environment* 2012b; **434**: 201-12.
474 Bowes MJ, Lehmann K, Jarvie H *et al*. Investigating biofilm biofilm response to changing phosphorus
475 concentrations in UK rivers using within-river flumes. In: Kirby C (ed.) *Role of Hydrology in Managing*
476 *Consequences of a Changing Global Environment* British Hydrological Society, 2010, 384-9.
477 Bray JR, Curtis JT. An ordination of the upland forest communities of Southern Wisconsin. . *Ecological*
478 *Monographies* 1957;**27**: 325-49.
479 Bruno, John F., John J. Stachowicz, and Mark D. Bertness. "Inclusion of facilitation into ecological theory."
480 *Trends in Ecology & Evolution* 2003,**18.3**: 119-125.
481 Burrell TK, O'Brien JM, Graham SE *et al*. Riparian Shading Mitigates Stream Eutrophication in Agricultural
482 Catchments. *Freshwater Science* 2013;**33**: 73-84.
483 Caporaso JG, Kuczynski J, Stombaugh J *et al*. QIIME allows analysis of high-throughput community sequencing
484 data. *Nature Methods* 2010;**7**: 335-6.
485 Chang YR. Bacterial community composition in stream biofilms is influenced by algal response to varying light
486 and phosphorus ratios*Natural Res & Env Sci* volume MSc. Urbana-Champaign: University of Illinois at
487 Urbana-Champaign, 2010.
488 Chistoserdova L, Jenkins C, Kalyuzhnaya MG *et al*. The enigmatic Planctomycetes may hold a key to the origins
489 of methanogenesis and methylotrophy. *Mol Biol Evol* 2004;**21**: 1234-41.
490 Chung S-M, Staub JE. The development and evaluation of consensus chloroplast primer pairs that possess highly
491 variable sequence regions in a diverse array of plant taxa. *TAG Theoretical and Applied Genetics*
492 2003;**107**: 757-67.
493 Clarke KR. Non-parametric multivariate analyses of changes in community structure. *Australian Journal of*
494 *Ecology* 1993;**18**: 117-43.
495 Cole JR, Wang Q, Cardenas E *et al*. The Ribosomal Database Project: improved alignments and new tools for
496 rRNA analysis. *Nucleic Acids Research* 2009;**37**: D141-D5.
497 Colijn F, Cadée GC. Is phytoplankton growth in the Wadden Sea light or nitrogen limited? *Journal of Sea*
498 *Research* 2003;**49**: 83-93.
499 Cross WF, Benstead JP, Frost PC *et al*. Ecological stoichiometry in freshwater benthic systems: recent progress
500 and perspectives. *Freshwater Biology* 2005;**50**: 1895-912.
501 Desortova B, Puncochar P. Variability of phytoplankton biomass in a lowland river: Response to climate
502 conditions. *Limnologica* 2011;**41**: 160-6.

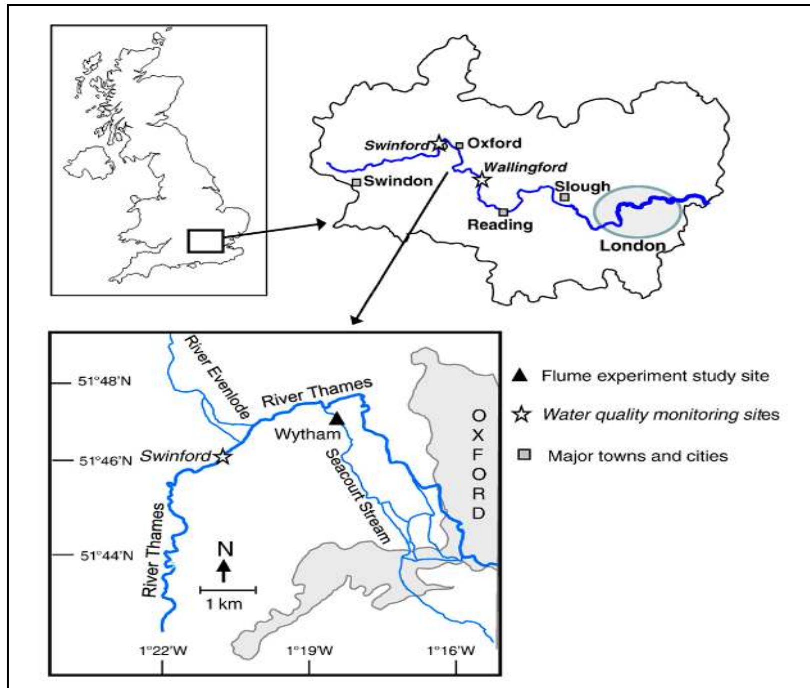
503 Dodds WK, Smith VH, Lohman K. Nitrogen and phosphorus relationships to benthic algal biomass in temperate
504 streams. *Can J Fish Aquat Sci* 2002;**59**: 865-74.
505 Dorigo U, Berard A, Rimet F *et al*. In situ assessment of biofilm recovery in a river contaminated by pesticides.
506 *Aquatic Toxicology* 2010;**98**: 396-406.
507 Dunfield PF, Yuryev A, Senin P *et al*. Methane oxidation by an extremely acidophilic bacterium of the phylum
508 Verrucomicrobia. *Nature* 2007;**450**: 879-U18.
509 EA. Water for life and livelihoods - River Basin Management Plan Thames River Basin District. Annex B: Water
510 body status objectives. Bristol: Environment Agency, 2009.
511 EA. What's in your backyard? Environment Agency, 2012.
512 Evans SG, Spillett PB, Colquhoun K. South-East Housing Development — The Quest For Sustainability: Water
513 And Sewerage Needs. *Water and Environment Journal* 2003;**17**: 257-61.
514 Ghermandi A, Vandenberghe V, Benedetti L *et al*. Model-based assessment of shading effect by riparian
515 vegetation on river water quality. *Ecol Eng* 2009;**35**: 92-104.
516 Guariento RD, Carneiro LS, Caliman A *et al*. How light and nutrients affect the relationship between autotrophic
517 and heterotrophic biomass in a tropical black water biofilm community. *Aquatic Ecology* 2011;**45**: 561-
518 9.
519 Hamady M, Knight R. Microbial community profiling for human microbiome projects: Tools, techniques, and
520 challenges. *Genome Research* 2009;**19**: 1141-52.
521 Hashidoko Y. Ecochemical studies of interrelationships between epiphytic bacteria and host plants via secondary
522 metabolites. *Biosci Biotech Bioch* 2005;**69**: 1427-41.
523 Hill WR, Fanta SE, Roberts BJ. Quantifying phosphorus and light effects in stream algae. *Limnol Oceanogr* 2009;
524 **54**: 368-80.
525 Hill WR, Knight AW. Nutrient and light limitation of algae in two northern California streams. *Journal of*
526 *Phycology* 1988: 125–32.
527 Hill WR, Roberts BJ, Francoeur SN *et al*. Resource synergy in stream biofilm communities. *Journal of Ecology*
528 2011;**99**: 454-63.
529 Hill WR, Ryon MG, Schilling EM. Light Limitation in a Stream Ecosystem: Responses by Primary Producers and
530 Consumers. *Ecology* 1995;**76**: 1297-309.
531 Hillebrand H, Sommer U. Response of epilithic microphytobenthos of the Western Baltic Sea to in situ
532 experiments with nutrient enrichment. *Marine Ecology-Progress Series* 1997;**160**: 35-46.
533 Huber JA, Morrison HG, Huse SM *et al*. Effect of PCR amplicon size on assessments of clone library microbial
534 diversity and community structure. *Environmental Microbiology* 2009;**11**: 1292-302.
535 Huse SM, Dethlefsen L, Huber JA *et al*. Exploring Microbial Diversity and Taxonomy Using SSU rRNA
536 Hypervariable Tag Sequencing. *Plos Genetics* 2008;**4**.
537 Huson DH, Auch AF, Qi J *et al*. MEGAN analysis of metagenomic data. *Genome research* 2007;**17**: 377-86.
538 Hutchins M, Johnson A, Deflandre-Vlandas A *et al*. Which offers more scope to suppress river phytoplankton
539 blooms: reducing nutrient pollution or riparian shading? *Science of the total environment* 2010;**408**:
540 5065-77.
541 Johnson AC, Acreman MC, Dunbar MJ *et al*. The British river of the future: How climate change and human
542 activity might affect two contrasting river ecosystems in England. *Science of The Total Environment*
543 2009;**407**: 4787-98.
544 Kelly MG, Wilson S. Effect of phosphorus stripping on water chemistry and diatom ecology in an eastern lowland
545 river. *Water Research* 2004;**38**: 1559-67.
546 Khare PK, Chaurasia JK. Seasonal variation of phytoplankton in four stations of Vijay Sagar Lake of Mahoba
547 (U.P.). *Journal of Environmental Research and Development* 2009;**4**: 123-31
548 Koren O, Knights D, Gonzalez A *et al*. A guide to enterotypes across the human body: Meta-analysis of microbial
549 community structures in human microbiome datasets. *PLoS computational biology* 2013;**9**: e1002863.
550 Kritzberg ES, Langenheder S, Lindstrom ES. Influence of dissolved organic matter source on lake
551 bacterioplankton structure and function - implications for seasonal dynamics of community composition.
552 *Fems Microbiology Ecology* 2006;**56**: 406-17.
553 Kruskal J. Nonmetric multidimensional scaling: A numerical method. *Psychometrika* 1964;**29**: 115-29.
554 Lenane R., Keeping Rivers Cool, Environment Agency, 2012
555 Lindemann SR, Moran JJ, Stegen JC *et al*. The epsomitic phototrophic microbial mat of Hot Lake, Washington:
556 community structural responses to seasonal cycling. *Frontiers in Microbiology* 2013;**4**: 323.
557 Mainstone CP, Parr W. Phosphorus in rivers - ecology and management. *Science of The Total Environment*
558 2002;**282-283**: 25-47.
559 Mccall, Stephanie, Michael Bowes, Tanya Warnaars, Michelle Hale, Jim Smith, Alan Warwick, and Cyril Barrett.
560 "Phosphorus enrichment of the oligotrophic River Rede (Northumberland, UK) has no effect on
561 periphyton growth rate." *Inland Waters* 4, no. 2 (2014): 121-132.
562 Mosisch TD, Bunn SE, Davies PM. The relative importance of shading and nutrients on algal production in
563 subtropical streams. *Freshwater Biology* 2001;**46**: 1269-78.
564 Neal C, Jarvie HP. Agriculture, community, river eutrophication and the Water Framework Directive.
565 *Hydrological Processes* 2005;**19**: 1895-901.
566 Neal C, Jarvie HP, Withers PJA *et al*. The strategic significance of wastewater sources to pollutant phosphorus
567 levels in English rivers and to environmental management for rural, agricultural and urban catchments.
568 *Science of The Total Environment* 2010;**408**: 1485-500.

569 Neal C, Neal M, Reynolds B *et al*. Silicon concentrations in UK surface waters. *Journal of hydrology* 2005;**304**:
570 75-93.
571 O'Driscoll JF, Harrison SSC, Giller PS. Do trees make a difference? An evaluation of the impact of riparian
572 vegetation on the ecology of nutrient poor headwater streams. *Polish Journal of Ecology* 2006;**54**: 695-
573 700.
574 Ohtsuka T. Epipelic diatoms blooming in Isahaya Tidal Flat in the Ariake Sea, Japan, before the drainage
575 following the Isahaya-Bay Reclamation Project. *Phycological Research* 2005;**53**: 138-48.
576 Paerl HW, Hall NS, Calandrino ES. Controlling harmful cyanobacterial blooms in a world experiencing
577 anthropogenic and climatic-induced change. *Science of The Total Environment* 2011;**409**: 1739-45.
578 Paithankar KR, Prasad KSN. Precipitation of DNA by polyethylene glycol and ethanol. *Nucleic Acids Research*
579 1991;**19**: 1346-.
580 Pillet L, de Vargas C, Pawlowski J. Molecular Identification of Sequestered Diatom Chloroplasts and
581 Kleptoplastidy in Foraminifera. *Protist* 2011;**162**: 394-404.
582 Quinn JM, Cooper AB, Stroud MJ *et al*. Shade effects on stream biofilm and invertebrates: An experiment in
583 streamside channels. *New Zealand Journal of Marine and Freshwater Research* 1997;**31**: 665-83.
584 Read, Daniel S., Michael J. Bowes, Lindsay K. Newbold, and Andrew S. Whiteley. "Weekly flow cytometric
585 analysis of riverine phytoplankton to determine seasonal bloom dynamics." *Environmental Science:*
586 *Processes & Impacts* 16, no. 3 (2014): 594-603.
587 Rier ST, Stevenson RJ. Effects of light, dissolved organic carbon, and inorganic nutrients on the relationship
588 between algae and heterotrophic bacteria in stream biofilm. *Hydrobiologia* 2002;**489**: 179-84.
589 Rier ST, Stevenson RJ, LaLiberte GD. Photo-Acclimation Response Of Benthic Stream Algae Across
590 Experimentally Manipulated Light Gradients: A Comparison Of Growth Rates And Net Primary
591 Productivity. *Journal of Phycology* 2006;**42**: 560-7.
592 Romanovskaya V, Titov A. Interrelation of monocultures in the microbe associations Methylococcus capsulatus
593 and Frateuria aurantiaca. *Mikrobiologichnyi Zhurnal* 1992;**54**: 3-8.
594 Rosemond A. Interactions among irradiance, nutrients, and herbivores constrain a stream algal community.
595 *Oecologia* 1993;**94**: 585-94.
596 Sanches LF, Guariento RD, Caliman A *et al*. Effects of nutrients and light on periphytic biomass and nutrient
597 stoichiometry in a tropical black-water aquatic ecosystem. *Hydrobiologia* 2011;**669**: 35-44.
598 Smayda TJ. Complexity in the eutrophication–harmful algal bloom relationship, with comment on the importance
599 of grazing. *Harmful Algae* 2008;**8**: 140-51.
600 Steinman AD. Does an increase in irradiance influence biofilm in a heavily-grazed woodland stream? *Oecologia*
601 1992;**91**: 163-70.
602 Triska FJ, Kennedy VC, Avanzino RJ *et al*. Effect of simulated canopy cover on regulation of nitrate uptake and
603 primary production by natural biofilm assemblages. In: III TDF, Bartell SM (eds.) *Dynamics of Lotic*
604 *Ecosystems* Ann Arbor, U.S.A.: Ann Arbor Science, 1983, 494 pp. .
605 Vanelslander B, Créach V, Vanormelingen P *et al*. Ecological differentiation between sympatric pseudocryptic
606 species in the estuarine benthic diatom Navicula phyllepta (Bacillariophyceae). *Journal of Phycology*
607 2009;**45**: 1278-89.
608 Warner, Richard E., and Kathleen M. Hendrix. *California riparian systems: ecology, conservation, and productive*
609 *management*. Univ of California Press, 1984.
610 Winterbourn MJ. Interactions among nutrients, algae and invertebrates in a New Zealand mountain stream. .
611 *Freshwater Biology* 1990: 463–74.
612 Ylla I, Borrego C, Romani AM *et al*. Availability of glucose and light modulates the structure and function of a
613 microbial biofilm. *FEMS Microbiology Ecology* 2009;**69**: 27-42.
614 Znachor P, Simek K, Nedoma J. Bacterial colonization of the freshwater planktonic diatom Fragilaria crotonensis.
615 *Aquatic Microbial Ecology* 2012;**66**: 87-94.
616

617 **List of Figures**
618
619 **Fig. 1**: Map of the Thames basin, showing the study location.
620 **Fig. 2**: Two mesocosms in the field. The upstream end is closest to the bottom edge of
621 the picture.
622
623 **Fig. 3:** Taxonomic distribution chart of all replicates at order level, based on relative
624 abundance of 16S OTUs from the pyrosequencing results. Algae did not get resolved
625 at order level. The most abundant organisms are listed in the legend. The x-axis
626 categories show the replicates: U1-U12 are the unshaded samples, S1-S12 the shaded
627 samples. Bacteria_Other defines the group of bacterial organisms that cannot be
628 identified to other taxonomic levels.
629
630 **Fig. 4:** Shannon diversity of both the algal and bacterial components of the biofilm
631 samples. ANOVA: $F = 32.70$ (Algae) and 9.168 (Bacteria), $p = 1.35e-05$ (Algae) and
632 0.007 (Bacteria)
633 **Fig. 5:** Log-transformed rank-abundance curves for both shaded and unshaded algae
634 and prokaryotes. Species that rank high in the Simper analysis are marked by name.
635
636 **Fig. 6:** NMDS of the shaded (S) and unshaded (U) replicates, divided into algae and
637 prokaryotes, from Seacourt Stream. Stress:0.19 (Algae)/ 0.18 (Bacteria).
638
639 **Fig. 7:** Networks of co-occurrence. Threshold: 0.5%, minimum/maximum:15%(2
640 samples)/100%(12 samples). Minimum probability that a co-occurrence between two
641 taxa can be observed: 95%.
642

643

644

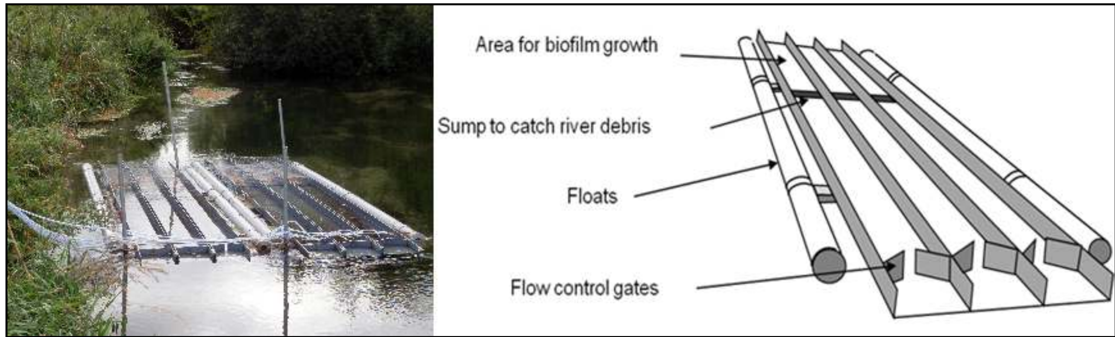


645

646 **Fig. 1:** Map of the Thames basin, showing the study location.

647

648



649 **Fig. 2:** design and photo of 6 flumes (facing downstream).

650

651

652

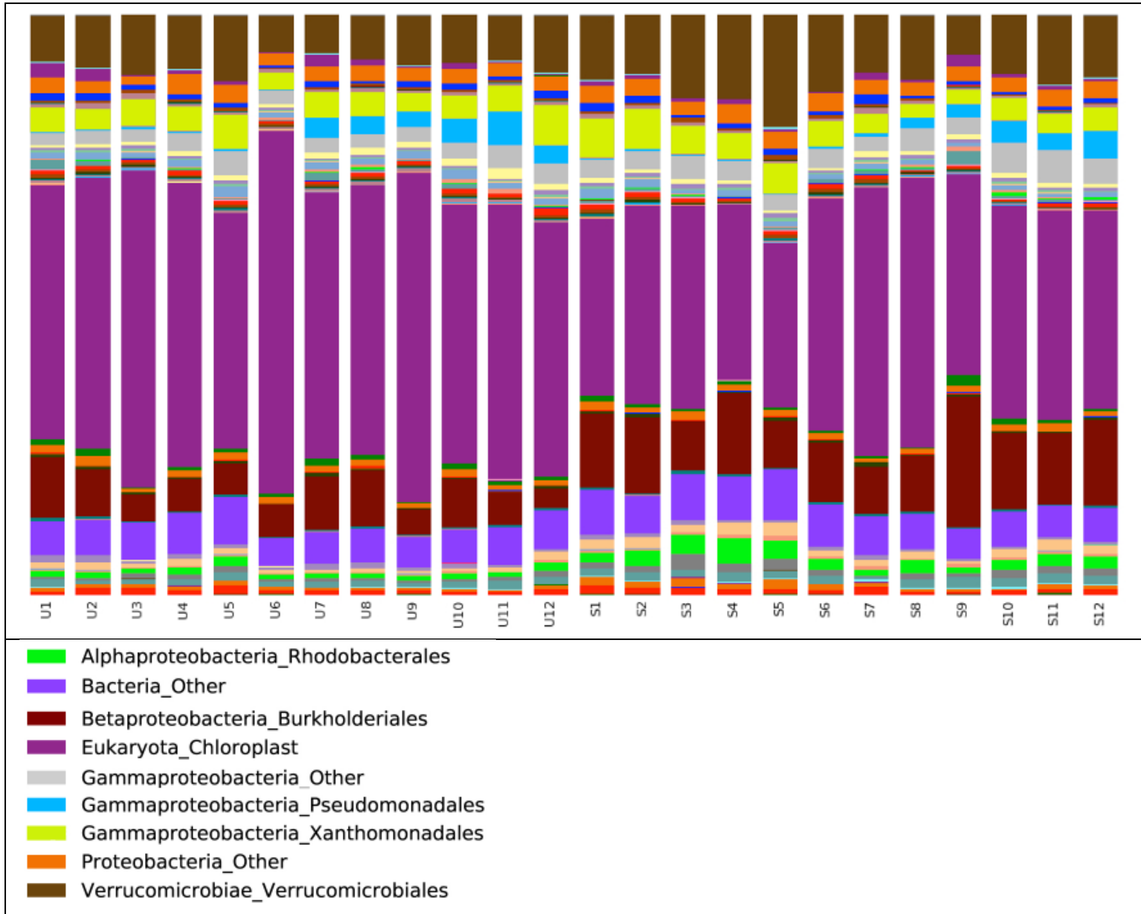
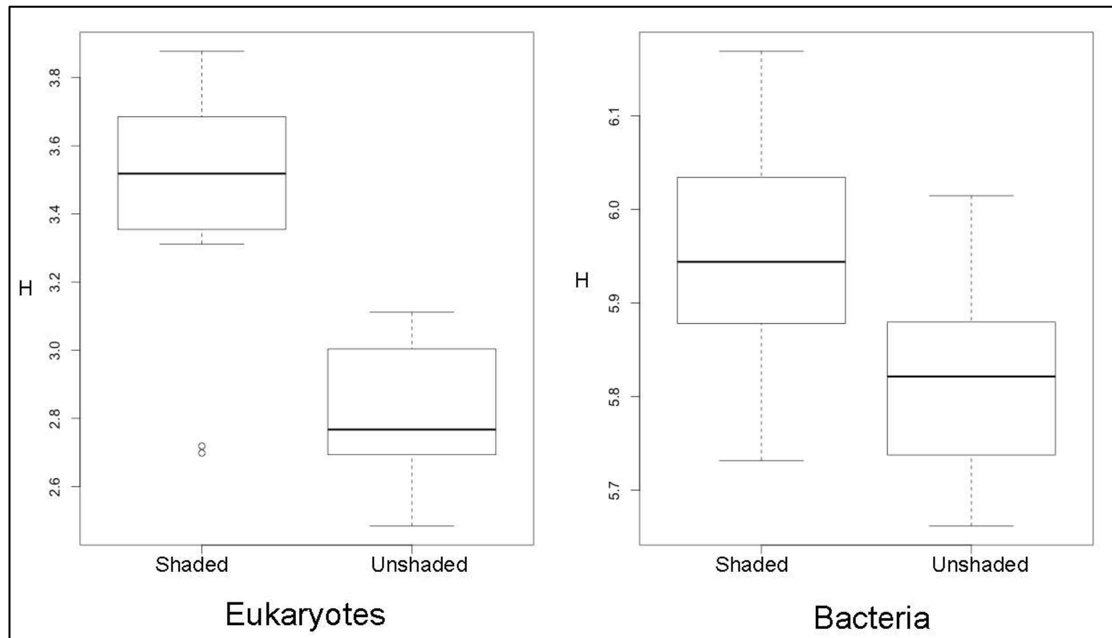


Fig. 3: Taxonomic distribution chart of all replicates at order level, based on relative abundance of 16S OTUs from the pyrosequencing results. Algae did not get resolved at order level. The most abundant organisms are listed in the legend. The x-axis categories show the replicates: U1-U12 are the unshaded samples, S1-S12 the shaded samples. Bacteria_Other defines the group of bacterial organisms that cannot be identified to other taxonomic levels.

659

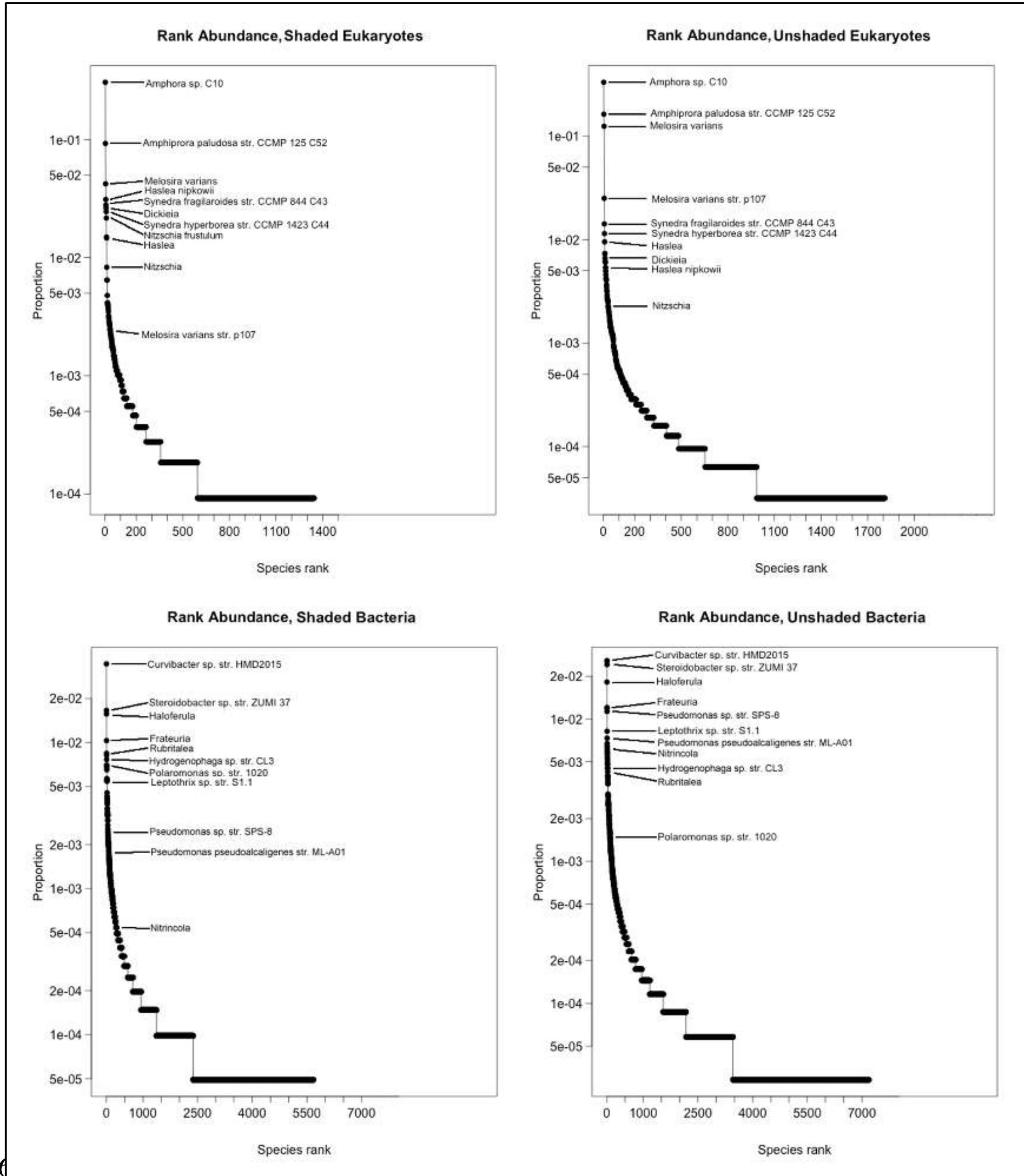


660
661 **Fig. 4:** Shannon diversity of both the algal and bacterial components of the biofilm samples. ANOVA: F = 32.70
662 (Algae) and 9.168 (Bacteria), p = 1.35e-05 (Algae) and 0.007 (Bacteria)

663

664

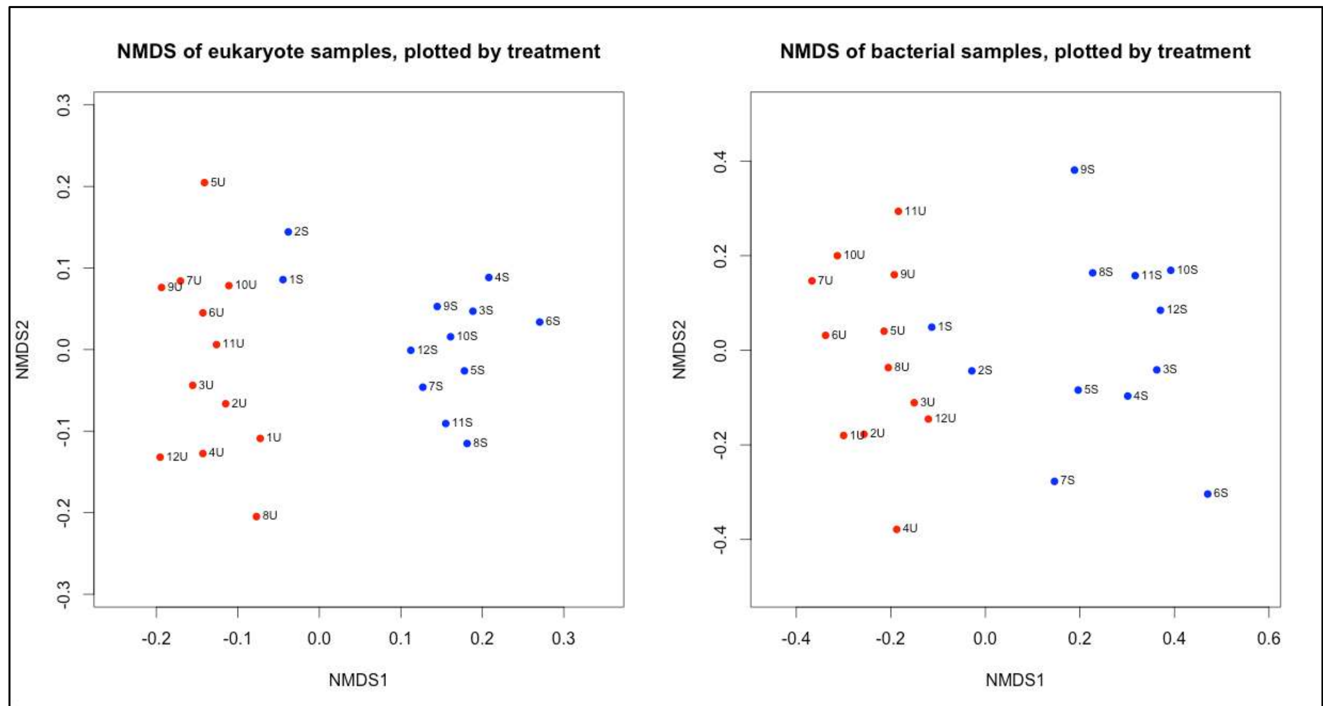
665



666

667 **Fig. 5:** Log-transformed rank-abundance curves for both shaded and unshaded Algae and Prokaryotes. Species that rank
668 high in the Simper analysis are marked by name.

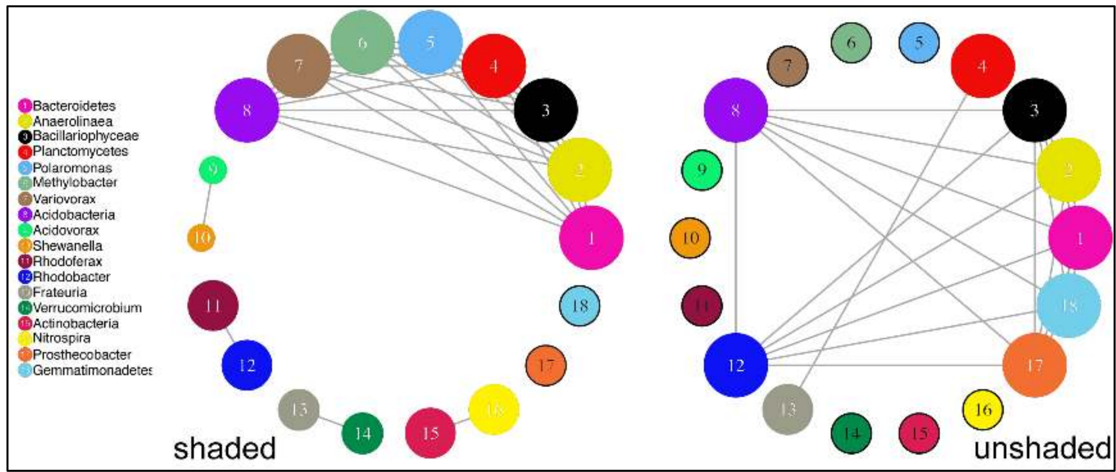
669



670
671 **Fig. 6:** NMDS of the shaded (S) and unshaded (U) replicates, divided into algae and prokaryotes, from Seacourt
672 Stream. Stress:0.19 (Algae)/ 0.18 (Bacteria).

673
674
675
676
677
678
679
680
681
682
683

684



685 **Fig. 7:** Networks of co-occurrence. Threshold: 0.5%, minimum/maximum:15%(2 samples)/100%(12 samples).
686 Minimum probability that a co-occurrence between two taxa can be observed: 95%. The numbers next to the nodes
687 specify in how many samples the taxa co-occur.