Treatment stage associated changes in cellular and molecular microbial markers during the production of drinking water at the Vansjø water works

Colin Charnock*, Eli Otterholt, Anne-Lise Nordlie

Faculty of Health Sciences, Oslo and Akershus University College of Applied Sciences, St. Olavs Plass, Oslo, Norway.

E-mail address: colin.charnock@hioa.no
*Corresponding author: Tel: 004767236275

ABSTRACT

The production of a drinking water that meets current aesthetic, microbiological and chemical standards, generally requires a combination of mechanical purification and disinfection in a multi-component treatment chain. Treatment choices and optimisation of water processing is best informed by using markers (including microbiological parameters) which indicate how each stage contributes to the production of the potable water. The present study combines culture-based and a number of culture-independent analyses to indicate what is happening at each stage of a state-of-the-art water treatment chain at Vansjø near the city of Moss in Norway. We show that particularly clarification with flotation and post-chlorination have profound and positive effects on water quality with respect to the removal and inactivation of microbes. Post-chlorination achieved better disinfection of the water than UV-treatment and was of paramount importance, as the penultimate step filtration through granular activated shed microbes to the water. Cloning and sequencing showed that some clones present in the raw water were detected at all stages in the treatment process, perhaps providing examples of microbes breaching physically all barriers in the treatment process. Results from the study should be useful in the improvement and maintenance of the treatment process at the Vansjø plant and others.

Keywords:

Water treatment efficacy
Removal of microbes
Total viable counts
Cloning
Bacterial community profiling

1. Introduction

In addition to its direct use as a thirst quencher, potable water is also used in the production of food, beverages and medicines. Treatment trains leading to the production of drinking water can show considerable variation, but most involve steps for the removal of particles (including microbial agents) and dissolved substances, and one or more disinfection stages to inactivate pathogens. Such an approach provides multiple barriers to water-borne microbes. Conventional treatment trains commonly involve a pretreatment stage (such as microstraining), followed by coagulation, flocculation and sedimentation/flotation (collectively called clarification). Filtration steps then remove particulate matter and in some cases (eg, granular activated carbon filters) colors, odors and tastes. Together, coagulation, flocculation and sedimentation can result in 1–2 log removals of bacteria, viruses and protozoa (Lechevallier, 2004). Less seems to be known about microbial removal based on clarification processes involving flotation. In dissolved air flotation, bubbles are produced by reducing pressure in a water stream saturated with air. The rising bubbles attach to floc particles, causing the agglomerate to float to the surface, where the material is skimmed off. Granular media filtration is widely used in drinking water treatment to remove flocculated matter and other particles. Granular activated carbon is used when both filtration of particles and adsorption of organic compounds are desired. For microbial pathogens, application of strong oxidizing compounds such as chlorine are commonly used for disinfection. Ultraviolet light inactivates microorganisms through reactions with microbial nucleic acids and most work has been on the effect of UV in the control of *Cryptosporidium* and other cyst-producing protozoa (Hinjen et al., 2006). Current drinking water directives in Europe and the US use culture-based techniques on the finished water to see if it meets prescribed quality standards.

However, 'the great plate count anomaly' (Staley and Konopka, 1985) indicates that culture-based techniques reveal only a small fraction of the microbes present. Our purpose with the present work is to see how each treatment stage impinges on the microbial content of a drinking water under production. To do so we have applied molecular and cellular analyses to study the fate of microbes on their passage through a multistage, state-of-the-art conventional drinking water treatment plant at Vansjø in the Southeast of Norway. Assessing the contribution of the individual stages in the treatment train to microbial removal and inactivation should be valuable in process optimisation.

Vansjø water works delivers yearly about 7.000.000 m^3 water to Moss and neighboring municipalities. The raw water for the production of potable water is drawn from lake Vansjø which is a relatively shallow body of water potentially exposed by its geography to both agricultural run-off and urban pollution. In order to meet the challenges posed by the poor quality of the water, the plant uses three major cleaning and 2 disinfection stages which are as follows. Abbreviations for the names of samples taken from each stage are given in parentheses. Step 1 is intake of the raw water (RW) from depth of 25m. Step 2 involves conditioning to increase the pH and a clarification stage including flotation (CF). Step 3 (2-media filtration, 2MF) uses 6 parallel filters consisting of layers of crushed stone, sand and leca. In step 4, the water runs to a basin from which it is pumped to the UV aggregate (before UV; BUV). Step 5 is UV-disinfection (after UV; AUV). Step 6 is granular activated carbon filtration using 6 parallel filters (GAC). Step 6 is post-chlorination with chloramine and pH-adjustment with $CaCO_3$, producing the drinking water (DW).

2. Materials and Methods

2.1 *Sampling*

Water samples were collected at the plant using aseptic techniques. Single samples (2l from each of the 7 stages of the treatment train) were collected in acid-cleaned bottles and maintained at 4 ºC prior to testing. All tests were initiated in the course of the same working day. Samples were collected in two rounds: one in October 2011 (Round 1, R1) and the other in December (R2). The treatment process is the same at the time of writing.

2.2 *Bacterial/Fungal heterotrophic plate counts and cellular identification*

2.2.1 *Bacteria*

Water samples (0.1 mL) were spread directly (drinking water) and also after dilution in sterile water (other samples) to give accurate plate counts on R2A agar (Oxoid™, Thermo Fisher Scientific, MA, USA). Plates were incubated aerobically at 22 ± 2 °C. Colony counts were made after 14 days (Reasoner and Geldreich, 1985). The mean of the colony counts or weighted mean (if more than one sample volume gave colony counts in the range 25-250) were recorded. Samples were also spread on sheep blood agar (Oxoid) and incubated aerobically at 37 ± 1 °C. Colony counts were made after 24 h and 48 h.

Bacteria were identified by partial sequencing of the 16S rRNA gene exactly as previously described (Otterholt and Charnock, 2011a) using the primers 341f and 907f and a small amount of colony material as template. Sequencing reactions were performed by a commercial laboratory (ABI-lab, University of Oslo, Oslo, Norway) using the PCR primers. Poor quality data and primer sequences were removed from the sequence terminals before similarity searches were performed.

155

156 **2.2.2 *Fungi***

157 Samples (10 and 250 mL) were filtered in a laminar flow clean bench using a filter manifold
158 and sterile filters (0.45 μm). Sabouraud dextrose medium (Oxoid) containing 100 μg mL^{-1}
159 chloramphenicol (Oxoid) was used as growth medium and plates were incubated for 72 h at
160 22 ± 2 °C. Distinct colony morphologies were examined in the microscope to confirm they
161 were fungal. The reported count includes only confirmed yeasts and molds. A forward primer
162 binding to 3′-end of 18S rDNA and a reverse primer binding to the 5′-end of 25/28S rDNA of
163 yeasts were used to amplify the intervening 5.8S gene and the noncoding Internal Transcribed
164 Spacer regions ITS1 and ITS2, exactly as previously described (Otterholt and Charnock,
165 2011a). Sequencing reactions were performed using the PCR primers.

166

167 **2.3 *Protozoa (culture and categorization)***

168 The presence/absence (P/A) of easily culturable protozoa in 250 mL water was investigated
169 exactly as previously described (Otterholt and Charnock, 2011b). Wet mounts from the plates
170 were examined in the microscope for P/A of amoebae and their cysts as well as ciliates and
171 flagellates

172

173 **2.4 *Live-dead cell staining***

174 In brief, 3 μL of solutions A and B from the staining kit (LIVE/DEAD *Bac*Light cell staining
175 kit; Molecular probes®, Thermo Fisher Scientific, MA, USA) were dispersed in 3 mL
176 particle-free sterile water ('the stain'). Water samples (10-100 mL) were filtered through a
177 0.22 μm black, polycarbonate filter. After filtration, 1 mL of the stain was added to the
178 mounted membrane. Thereafter the membrane in its manifold was kept in the dark for 20 min.
179 For purposes of microscopy, 1 drop of *Bac*Light mounting oil was added directly to the

membrane and covered with a coverslip. One drop of Citfluor™ anti-fadent (CitiFluor Ltd, London, UK) was added to the coverslip and the membrane was examined by fluorescence microscopy (Olympus BX40 , GmbH). On each filter no fewer than 400 cells and usually >1000 were counted in 8-15 fields.

2.5 *Biolog® GN2* (Biolog, Inc., Hayward, Calif, USA)

Water samples (0.15 mL) from each treatment stage were pipetted into each of the 96 wells (1 control and 95 different carbon substrates). During incubation microbes oxidize substrates in the plate wells and simultaneously reduce the colourless tetrazolium dye to a violet formazan. Wells were scored for substrate oxidation after incubation for 14 days at 22 ± 2 ºC based on the fulfillment of 2 criteria: a visible violet coloration in the wells and an optical density (OD_{595nm}) 30% or more above the control well value. Plates were read using a Victor Multilabel Counter (Perkin Elmer, Turku, Finland)

2.6 *Clone libraries (R2 of testing only).*

Clone libraries were generated using primers 341f and 907r (see 2.2.1) and template DNA derived from the bacterial communities at each treatment stage. DNA was isolated as follows: 500 mL of water sample was filtered through a 0.2 µm GTP isopore track-etched polycarbonate filter (Merck Millipore, Darmstadt, Germany). For the isolation of DNA from membranes the PowerWater® kit was used (Mo Bio, CA, USA). In brief, the filter was rolled and inserted into the bead tube provided in the kit. Thereafter, 1mL of PW1 from the kit was added and the 'alternative method' for DNA isolation was followed exactly as described. Eluted DNA was quantified and the 260/280 ratio measured using a nanodrop device (Saveen Werner, Sweden). DNA was stored at -20 ºC until required. Cloning and transformation was performed using a TOPO® TA Cloning Kit (Invitrogen™, Thermo Fisher Scientific). In brief, about 30 ng of PCR product per µl (10 ng) vector were ligated as specified in the kit's

207 *complete protocol*. PCR amplicons for ligation were generated using the primer set 341f/907r
208 exactly as described for colony-PCR (see, section 2.2.1) except that about 20 ng biofilm DNA
209 per 50 µl PCR-reaction was used as template in the reaction.
210
211 **2.7** ***RNA isolation from microbial communities and production of cDNA from total RNA***
212 ***(R2 of testing only).***
213 For the isolation of total RNA, 500 mL of water sample was filtered through a 0.2 µm GTP
214 isopore track-etched polycarbonate filter (Millipore-Merck). The RNA Power water kit (Mo
215 Bio) was used for the extraction and purification of total RNA from filters. In brief, the filter
216 was rolled and inserted into a kit bead tube, thereafter the supplied protocol was followed.
217 After elution, the RNA concentration was measured using a nanodrop (see 2.6). RNA was
218 stored at -20 ºC until required.
219 For the production of cDNA the ominiscript RT system (Qiagen, Germany) was used.
220 All equipment was bought or prepared to be RNase free. The RT mastermix (including
221 RiboGuard RNase inhibitor; Epicentre®, Wisconsin, USA) was made following the kit
222 instructions. The template was 2 µL RNA isolated from water treatment steps 1-7 (variously
223 8 – 20 ng RNA). Reaction mixes contained random nonamers (obtained from Eurofins
224 Scientific) as primers. The reaction mix was incubated at 37 ºC for 90 min followed by 93 ºC
225 for 5 min with a final rapid cooling on ice. The resulting cDNA was store at – 20 ºC prior to
226 use in RAP-PCR (see 2.8).
227
228 **2.8** ***AP-PCR, RAP-PCR and bacterial community profiling***
229 In order to generate community profiles, the M13 reverse primer (5'-CAG GAA ACA GCT
230 ATG AC-3') was used (Eurofins scientific). The reaction mix (50 µL) consisted of 10 µL
231 PCR hot start buffer (Promega, Wisconsin, USA) 3 µM primer, dNTPs at 0.1 mM, 2 mM

232 $MgCl_2$, 0.75 U of hot-start polymerase (Promega), 1-5 ng DNA (AP-PCR) or cDNA (RAP-
233 PCR), PCR-grade water to a final volume of 50 µL. PCR conditions were as follows : 2
234 cycles of 95 ºC/5 min; 37 ºC/ 5 min; 72 ºC/5 min. Followed by 45 cycles of 95 ºC/1min; 37
235 ºC/1min; 72 ºC/2 min. With a final extension of 72 ºC/10 min. Approximately 100-200 ng of
236 DNA from each reaction was electrophoresed in an 8% acrylamide gel using a 1 x TAE buffer
237 system at 40V for about 20h (or until good separation of the bands was obtained). Gels were
238 silver stained (Silver Staining Kit, Fisher Scientific). For purposes of comparison a DNA
239 marker was included (DNA ladder 1 kb+; Invitrogen, Life Technologies). Gels were
240 visualized using the Doc™ XR + System. Gel analysis and construction of a phylogenetic tree
241 based on band matching was performed using the Quantity-1 (Bio-Rad) software package.
242 Phylogenetic similarity trees were constructed based on calculated Dice coefficients of
243 correlation and using the unweighted pair group method with arithmetic averages clustering
244 (UWPGMA).

245
246 **2.9 *Sequence similarity searches and statistical analyses***

247 Sequences were compared for similarity with previously published ones using BLAST [basic
248 local alignment search tool] (Altschul et al*.,* 1990). Sequences were also analysed using the
249 'Classifier' (Wang et al., 2007) function with the'16S rRNA training set' and the 'Warcup
250 ITS training set for fungi'. Both are available at the ribosomal database (RDP) site (Cole et
251 al., 2013). For the production of pie charts, the RDP Classifier was used to assign sequences
252 to taxa using an 80% confidence level. Phylum- and class-level (for the *Proteobacteria*)
253 designations were used. The RDP 'library compare' tool was used to investigate statistical
254 differences between clone libraries. With this, sequences are first assigned to taxa (80%
255 confidence level) and then *P* values are estimated for pairwise comparisons of the same taxon
256 (number of sequences) in each library.

Sequences were binned into OTUs at the 97% and 98% similarity levels across the whole sequenced length using BLASTclust at http://toolkit.tuebingen.mpg.de/blastclust. The number and abundance of OTUs were used to evaluate classical non-parametric diversity, richness and evenness indices. The observed coverage richness (an assessment of the number of species present), in clone libraries was calculated using the Chao I indicator (Chao, 1984). Species richness and diversity were estimated using the Margalef and Shannon-Wiener indices. The latter is much used and takes into account both the number and evenness of species and is more sensitive to changes in abundance of the rarer groups (Hill et al., 2003). The Berger-Parker dominance index, which is a simple measure of the importance of the number of clones in the most abundant OTU was also applied. Coverage was estimated using Good's index (1953).

3. Results and Discussion

3.1 *Bacterial heterotrophic plate counts and identification of dominant species*

Table 1 provides an overview of the results from 2 rounds of testing. Numbers in parentheses are the results from a previously published round of testing (Charnock et al., 2010). The most important result in terms of consumer health is that no microbes (bacteria, fungi or protozoa) were ever grown from the DW samples at the plant. Thus the treatment train in its entirety produces a DW of high hygienic quality with respect to its content of easily cultured microbes. However, measures of microbial numbers after each treatment show that post-chlorination is essential to achieving this result (Table 1). This is particularly so as UV-treatment was only moderately effective in the inactivation of bacteria and the penultimate step, GAC filtration, increased colony counts (Table 1 and Fig. 1). Although significantly less effective than post-chlorination it is likely that the gross reduction in the R2A plate count attributable to UV-treatment facilitates the effects of post-chlorination. R2A agar was chosen to give an indication of changes in the total heterotrophic plate count as this medium was designed with the intention of detecting greater numbers of bacteria indigenous to water-supplies (Reasoner and Geldreich, 1985). Increases in colony counts in water passing through GAC is possibly a consequence of the shedding of bacteria from putative biofilms on the filtration material. Similar effects have been reported and discussed previously (Hammes et al., 2008). Rapid growth on blood agar at 37 ºC is an ability common to many pathogens of man. Table 1 shows that CF effectively removed the blood-agar count and that subsequent treatment steps did not restore it. In 2 of 3 investigations CF increased the R2A plate count. In this regard it may be relevant that the mainly pigmented colonies grown from the RW,

were generally replaced by off-white colonies after CF. The significance of this is not known, but the change shows that this step has a characteristic effect on the microbial community. The results could be explained by removal of the pigmented fraction and the addition of non-pigmented species, or by CF stimulating changes in cell physiology leading to a loss of pigment formation. This theme is returned to (see 3.6.2) when the clone libraries from each stage are compared. However, the increase in R2Acolony counts after CF seen on two occasions is at odds with all other measures (see Table 1, as well as live-dead staining 3.4 and GN2-testing 3.5) which indicate that CF significantly improves the microbiological quality. R2A is a relatively nutrient poor medium compared with blood agar. However, in our previous study (Charnock et al., 2010), 'water agar', which contains no added nutrient, showed that CF reduced colony counts by about 87%. We have been unable to find comparable studies in the literature, and thus there is a need for further investigation. Fig. 1 charts the characteristic changes in the R2A plate count.

In order to gauge to what extent the results for colony counting might represent specific health issues, some of the colonies were identified. The GAC filtration step, in both the present and previous (Charnock et al., 2010) study was shown to increase the R2A plate count, suggesting contamination of the material with microbes (Table 1; Fig. 1). Microbial communities in GAC filters can be useful in the removal of the biodegradable part of organic matter, and thus help to control microbial regrowth in the distribution system (Kaarela et al., 2015). The GAC plate counts were dominated by a single, flat, greenish colony morphology. Due its dominance and persistence, the strain was identified by the PCR-sequencing approach. RDP classifier confirmed the identification to the genus level as *Flavobacterium.* The genus *Flavobacterium* (phylum, *Bacteroidetes*) is widespread in nature and isolated from many freshwater and soil habitats (McCammon and Bowman, 2000). *Bacteroidetes* were also detected in the clone library from the GAC and other stages (see 3.6.2). The isolate might

form biofilms on the GAC. A putative ability to survive low nutrient stress could also be a factor for survival and proliferation in the GAC filters. It will be interesting to see if future work reveals that *Flavobacterium* spp. are common inhabitants of GAC-filters and if so, if they contribute to the removal of assimilable organic carbon from the water.

3.2 *Fungal heterotrophic plate counts and identification of dominant species*

The heterotrophic plate counts on Sabouraud agar suggest that CF, UV-treatment and post-chlorination improve water quality with respect to its content of fungi . In one instance, 2MF shed large numbers of fungi to the water, which were not adequately removed by UV-treatment, representing a breach of one of the disinfection stages (Table 1). This is likely caused by intermittent colonization of the filtration mass with fungi when the filtration time becomes too long, or if the filters are not adequately maintained. This effect has also been observed previously at the plant at the GAC stage (Charnock et al., 2010). An attempt to address this issue was the implementation of more regular flushing of the GAC with an interval reduction from 200 h to 150 h (Sundby pers. comm.).

Fungal colonies from individual treatment stages in the present study were identified by the PCR-sequencing approach. Many fungi grew from the RW (Table 1). Four colonies from each round of testing were identified by the sequencing approach. Using the RDP Classifier with the 'Warcup ITS training set for fungi' database for sequence comparisons, the following identifications were made for RW: Round 1: Saccharomycetales (two identical sequences), Nakazawaea (possibly *holstii*) and *Cryptococcus* (possibly *magnus*). Round 2: *Candida* (possibly *sake*) - 4 identical sequences. The use of 'possibly' here and elsewhere is used to indicate that the RDP-classification (see materials and methods) produced a boot strap value of 1.0 at the species level. Thus, as in our previous investigation (Charnock et al., 2010), *Candida sake* (putative) was shown to be among the easily cultured yeast in the lake

369 water. Candida sake infections are rare, but the species has been shown to cause severe
370 infections including fungal endocarditis, peritonitis and bloodstream infections (Juneja et al.,
371 2011). Three colonies from each round of testing were identified to gain information on the
372 fungi emerging from the CF stage. Round 1: two colonies were *Cryptococcus* (possibly sp
373 AL_V), the third was *Cryptococcus* (possibly *magnus*). Round 2: gave three colonies of
374 *Cryptococcus* (possibly *magnus*) – 3 identical sequences. With respect to samples taken after
375 2MF, yeasts of a single white morphology made up the whole fungal plate count in R1 of
376 testing and pink colonies made up almost the entire plate count in R2. Sequencing studies
377 showed the former to be *Cryptococcus* (possibly *magnus*). Four colonies analyzed after UV
378 treatment were also shown to be the same *Cryptococcus* (possibly *magnus*) clone. The pink
379 colonies (CF, R2) were identified as *Rhodotorula* (possibly *larynges*). In our previous
380 investigation at the plant (Charnock et al., 2010), a putative *C. magnus* clone was found in
381 high numbers in water emerging from both the 2MF and GAC filters. Thus *C. magnus/R.*
382 *laryngis* may colonize intermittently the filtration masses. Furthermore, when present in high
383 numbers, *C. magnus* was only partially inactivated by post-chlorination (Table 1). *C. magnus*
384 has been isolated from nasal cavities of pediatric cancer patients with leukemia (Khan et al.,
385 2011). However, there is no indication in the literature that *C. magnus* is an important cause
386 of cryptococcosis in humans. Although not regarded as a pathogen of humans, a later
387 emergence as an opportunistic pathogen of man cannot be ruled out. It will be interesting to
388 see if future studies reveal the presence of this species in filtration masses.
389
390 **3.3** ***Protozoa (culture and characterization)***
391 A wide variety of amoebae, ciliates and flagellates were grown from the raw water in both
392 rounds of testing. Trophozoites and cysts typical of *Hartmannella* and *Acanthamoeba* were
393 visible. Subsequently only one sample, (GAC; round 1), gave growth of protozoa: in this case

394 a single amoebal type resembling *Acanthamoeba* was found. The data is insufficient to draw
395 conclusions other than that the treatment train reduces protozoal contamination, and that no
396 protozoa in the present or previous (Charnock et al., 2010) study have as yet been found in the
397 DW. The presence of a putative *Acanthamoeba* after the UV treatment suggests that some
398 cysts escape inactivation. However, the method employed is a simple presence/absence test
399 and is insufficient to provide %-reductions in amoebal content. That the rich population of
400 ciliates and flagellates seen in the raw water is effectively removed already at the stage of CF
401 provides further support for the importance of this step in microbial control.
402
403 **3.4 *Live-dead cell staining***
404
405 In order to provide a truer evaluation of treatment efficacy (ie, the removal and inactivation of
406 microbes), live-dead cell staining was performed. The sum of membrane-intact ('live') and
407 membrane-damaged ('dead') cells has been shown to give comparable results to total cell
408 counts using acridine orange (Boulos et al., 1999). The additional advantage of identifying the
409 living fraction is that this approach enables an evaluation of the disinfection efficacy of UV
410 and chlorination. The treatment train in total gave between 2.76 – 2.94 $\log_{10}$-reductions in the
411 live cell count, which is similar to the value (2.81 – 2.89) reported for a conventional
412 coagulation-filtration train with post chlorination (Hoefel et al., 2005). Figure 2 shows
413 changes in the total cell count (TCC) and live cell count (LCC) along the treatment train. The
414 TCC count in the RW was variously ~ 3.5 – 6.0 x 10^6 mL^{-1} (Fig. 2). Approximately 30% of
415 the bacteria making up the TCC were metabolically active. Thus, for example in R2 of testing
416 only about 0.1% of the metabolically active fraction grew on R2A agar (Table 1) and
417 treatment efficacies based on total heterotrophic plate counts must be interpreted with this
418 limitation in mind. These values are similar to those reported by Hoefel et al (2005) where
419 the metabolically active fraction of a reservoir water made up 45-50% of the total count and

where the plate count was 0.03 – 0.07% of the total count. Ćirić et al., (2011) reported a raw reservoir water TCC of about 1.4 x 10^6 mL^{-1}. Similarly, a lake water in Israel showed also TCC in the order of ~2 x 10^6 - 1 x 10^7 mL^{-1}, of which about 8% were membrane-intact (Berman et al., 2001). Both values are similar to that of lake Vansjø in the present study. The TCC counts show that CF removed the bulk (of the order of 70-80%) of both TCC and LCC and was the treatment stage giving the greatest total removal. Notably, UV treatment gave reductions of about 40% in the metabolically active fraction entering the treatment unit (BUV), whereas chlorination achieved about a 99% reduction of the LCC emerging from the GAC filters. These results are in keeping with the trends in colony count data (Table 1) and show the extreme efficacy of the post-chlorination stage. The shedding of bacteria from the GAC filter to the water column seen in the colony count data, is also seen as a peak in the TCC and LCC curves (Fig. 2). Increases in cell numbers in the liquid fase of GAC filters has been reported previously (Hammes et al., 2008). In that case, GAC filters were, as in the present study, the penultimate treatment stage and cells were efficiently removed by the final ultrafiltration stage.

For DW, the TC was 3.4 – 4.5 × 10^5 mL^{-1} (R1-R2) and the LCC was 2 × 10^3 mL^{-1} (both rounds). In samples of tap water measured in the same manner as in the present study (ie, using live-dead cell staining), a TCC value of ~ 1 x 10^5 mL^{-1} was obtained (Boulos et al., 1999). Kahlisch et al., (2012) report tap water concentrations of ~ 4-5 × 10^5 mL^{-1}, of which about half of the cells were membrane intact and half were membrane damaged. The TCC are in these instances similar to the present study. However, the number of metabolically active cells in the drinking water sample taken at the Vansjø plant (~ 2 x 10^3 mL^{-1}) was much lower than that reported by Kahlisch et al., (2012) for tap water. The most likely explanation of this phenomenon is that the sample in the present study was taken at the water works directly after chlorination and before distribution. That colony counts typically increase during distribution,

445 and often dramatically when the chlorine residual diminishes and water comes into contact
446 with pipe biofilms is common knowledge, and has also been reported for the area supplied by
447 Vansjø plant (Charnock et al., 2010).
448
449 **3.5** ***GN2 analyses***
450 GN2-plates have been used in structure analysis of complex microbial populations in natural
451 samples (O'Connell et al., 2002; Stefanowicz, 2006). When applying a mixed population of
452 microbes to each well, the result will reflect the metabolic potential of the microbial
453 population, ie, both cell numbers and types and their combined potential to oxidize the
454 substrates. Table 1 shows that CF gave a reduction of 30-54% in the number of substrates
455 utilized relative to the RW. Again, supporting the proposed great effect of the CF stage on the
456 microbiological quality of the water. Subsequent stages prior to UV-treatment showed
457 variable results. Both disinfection stages significantly decreased the combined metabolic
458 potential of the microbial populations, with post-chlorination completely eliminating all
459 metabolic activity as measured by the GN2 system. Metabolic fingerprinting of the water thus
460 supports the main trends seen in colony counts (Table 1) and live-dead cell staining (Fig. 2).
461
462 **3.6** ***DNA extraction and cloning and sequencing***
463 **3.6.1** ***Binning of sequences into OTUs and treatment stage associated changes in clone***
464 ***libraries***
465 For R2 of testing clone libraries were also produced. The yield of total community DNA from
466 filters was variously 5-35 ng μL^{-1} of good quality (260/280: a ratio of ~1.8 DNA). This is
467 within the proposed expected range for the kit. After cloning and transformation, 50 colonies
468 from each treatment stage were randomly chosen for sequencing. In a number of instances,
469 clones revealed truncated or missing sequences. The exact number of clones from which

complete sequences (ie, exactly within the binding sites for each primer) are given in table 2. Using the BLASTclust (release 2.18.1) programme, sequences were binned into OTUs using the percentage identity threshold function. These and the outcome of non-parametric treatment of the data to reveal coverage, richness and evenness are summarized in Table 2.

The Margalef and the Shannnon-Wiener indices were higher for RW than DW whereas the Berger-Parker index was lowest in the RW. This indicates that the treatment train in its entirety reduces both species richness and diversity. The Chao 1 index suggests that only ~ 20 - 25% of the predicted number of OTUs present in the total sample pool are represented in the clone libraries. Whereas Good's coverage estimator indicated that the relative sampling coverage was ~ 73- 76%. These results taken together indicate that a greater number of clone sequences would be required for the libraries to give a good representation of the changes occurring. Cloning usually under samples true diversity in a complex ecosystem, retrieving only dominant taxa (Godoy-Vitorino et al., 2013). However, some salient details on clone library compositions (ie, taxonomy) emerge from the data set as presented in section 3.6.2 below.

3.6.2 *Taxonomic hierarchy of the clones and treatment-stage associated changes*

Taxonomic assignments were made at the phylum and, for the Proteobacteria, class levels using the RDP classifier (Fig. 3). The clone libraries from each stage are visually similar. Alpha- and beta-proteobacteria and actinobacteria dominate the libraries from every stage. These taxa are common in freshwaters (eg Burkert et al., 2003; Nielsen et al., 2006) and thus their dominance may not be surprising. However, the results suggest that no stage in the treatment causes a radical restructuring of the bacterial content. This is in contrast to for example the contamination of the water with the fungus *C. magnus* after 2MF (Table 1). It has previously been reported (based on 16S pyrosequencing) that the raw water seeded the

drinking water microbiome (Pinto et al., 2012), and this perhaps explains the general similarities in the bacterial clone libraries taken at each stage in the present study. However, the referenced study also clearly shows that water treatment (particularly dual-media sand filtration) shapes the drinking water bacterial community. In order to provide more detailed taxonomic comparisons, the RDP 'library compare' function was used to investigate statistical differences between individual taxa in the libraries. The tool generates *P* values for pairwise comparisons. As mentioned above, the R2A plates for the raw water and after CF samples showed a shift from pigmented to non-pigmented colonies. However, the clone libraries are basically similar (Fig. 3a and b), with the raw water containing only a few clones of planctomycetes and firmicutes not found in the flotation fraction. There is a significant difference between the RW and CF libraries ($P < 0.05$) with respect to their content of Actinobacteria, but not for Proteobacteria. It is possible that the Actinobacteria (which make up a larger portion of the CF sample) are generally less pigmented than other groups in the samples, explaining the visual difference in the R2A plate cultures. There were significantly more Alphaproteobacteria in the water after 2MF than in the water emerging from the CF stage ($P < 0.01$), but no significant difference at the phylum (Proteobacteria) level. The majority of Alphaproteobacteria in both samples scored as belonging to the SAR 11 clade (possibly *Candidatus Pelagibacter* sp.). However, this taxon (which is specifically discussed in 3.6.3) was present throughout the treatment train, and could not be traced directly to the filtration mass. The clone libraries for AUV and GAC stages were compared to see if the filter material causes changes in the microbial composition. However, no significant differences at the phylum and class levels were detected. This suggests that the increase in bacterial numbers after GAC, is owing to the re-introduction of bacteria already common at other stages of the treatment process. In a study of GAC filters used in the treatment of lake water, it was reported that bacterial communities were diverse but dominated by α-

520 Proteobacteria, β-Proteobacteria, and Nitrospira (Kaarela et al., 2015). Nitrospira was not
521 found in the present study. However the former named taxa made up 60% of the GAC clone
522 library (Fig. 3f). There were significantly more Proteobacteria and *Candidatus Pelagibacter*
523 sp. Clones ($P < 0.05$) in the DW than in the RW. This suggests these groups may be enriched
524 during the treatment process. Sequencing of a greater number of clones (as discussed over)
525 would provide a more robust data set, and other inter-library differences might then emerge.
526

527 **3.6.3** ***Clones found at all stages of the treatment train***

528 At 100% sequence similarity, 37 clones of 521 bp, (13% of the total of 279 clones sequenced)
529 clustered as a single OTU. Furthermore, at least two clones from each stage of the water
530 treatment are represented in this OTU. This OTU thus represents a major taxon which is
531 ubiquitous throughout the treatment train, and which probably originates in the raw water and
532 remains throughout the treatment process, breaching the barriers at each stage. When
533 restricting the search to named genera, the BLAST '16S ribosomal RNA' database gave a
534 single best hit of 92% similarity over the whole sequence length with *Candidatus*
535 *Pelagibacter* strain sp. IMCC9063 strain IMCC 9063. The identification *Candidatus*
536 *Pelagibacter* sp. was obtained using the RDP Classifier. The classifier contains only a single
537 *Candidatus Pelagibacter* sequence which is also that of strain IMCC 9063. The finding of this
538 genus is discussed below. At 98% similarity, a total of 21 clones from all treatment stages
539 formed a second OTU. Nine of the clones continued to cluster at the 100% level and these
540 were searched for similarities. The best hit using the BLAST '16S ribosomal RNA' database
541 was 91% similarity over 99% of the sequenced length with *Demequinia sediminicola*,
542 *Demequina aestuarii* and *Streptomyces aomiensis*. At an 80% confidence level the RDP
543 classifier assigned the sequence to the order Actinomycetales. Actinomycetales are an order in
544 the phylum Actinobacteria which makes up a significant proportion of all the clone libraries

545 (Fig. 3). Given the uncertainty in the identification at the genus level, the clone was not
546 considered further.

547 *Candidatus Pelagibacter* belongs to the SAR11 clade of alpha-proteobacteria. The
548 clade is abundant, globally distributed and contains the most numerous heterotrophs in marine
549 euphotic zones (Morris et al., 2002). *Candidatus Pelagibacter ubique* which was the first
550 member of the clade to be isolated and is by far the most studied (Morris et al., 2002)
551 possesses many unusual features for a free living organism, including an extremely small,
552 streamlined genome which lacks many genes and pathways that are otherwise common in
553 bacteria (Giovannoni et al., 2005). *Candidatus Pelagibacter* sp. IMCC 9063 strain IMCC
554 9063 (Oh et al., 2011) was isolated from surface sea water near Svalbard and is only distantly
555 related to other cultured representatives of the clade. At present no phenotype-based
556 description of the strain is available. A freshwater clade of the SAR11 has been identified and
557 is referred to as LD12 (Bahr et al., 1996). Given the abundance of the SAR11 clade in marine
558 and freshwaters, its presence throughout the treatment chain may not be surprising. If not
559 present solely by virtue of their general abundance, these oligotrophic, metabolically
560 restricted ultramicrobacteria may perhaps be suited to passage through treatment system
561 barriers.

562

563 **3.7** ***The yield of community RNA. AP-PCR and RAP-PCR profiling***

564 RNA extraction from filters gave yields of ~ 4-10 ng μL^{-1} RNA which was used in the
565 production of cDNA for subsequent RAP-PCR. Figure 4 shows the resolution of DNA-
566 fragments obtained after RAP-PCR (lanes 1-7) and AP-PCR (lanes 9-15) from each treatment
567 stage. The number of bands in each of the profiles was also estimated using the detect band
568 function in the Quantity-1 (Bio-Rad) software package. Figure 5 shows the changes in band

numbers for both RP-PCR and RAP-PCR as percentages of the highest number of bands in each profile.

Shifts in microbial communities along the treatment train are evident based on the general appearance (Fig. 4), corresponding similarity dendrogram (Fig. 6) and numbers of bands in each profile (Fig. 5). Band numbers as well as sizes is an index which has been used by others to express diversity and bacterial burden (Sharma et al., 2004, 2006). The AP-PCR profiles show a reduction in this index through the first 5 stages of treatment, indicating a general removal of genomic DNA from the water. The AP-PCR curve follows closely that of TCC (Fig. 2). The increase in the number of bands after GAC may indicate the general deterioration in the microbiological quality discussed extensively above. Fig. 4 shows a clear visual similarity between several of the individual AP-PCR profiles. Clustering analyses (Fig. 6) reflects this. Samples after and including 2MF form a cluster at about 50% similarity. The result suggests that 2MF and later stages have less influence on the bacterial content of the water than CF. Also the TCC plot levels out after CF (Fig. 2).

RAP-PCR is based on arbitrarily-primed reverse transcription of total community RNA using a random primer (here M13). This is followed by an arbitrarily-primed PCR which uses M13-generated cDNAs as template, and finally gel electrophoresis of the products. We have applied RAP-PCR as it has been suggested that the technique can be used to investigate the metabolic profile of microbial communities (Aneja et al., 2004; Sharma et al., 2004, 2006; Demanou et al., 2006). The RAP-PCR profiles, however, are not easily explicable in terms of proposed underlying theory, with only the salient increase after GAC and subsequent decrease after post-chlorination indicating that bands numbers might reflect total metabolic activity. One cluster (Fig. 6) consists of the all of the cDNA samples, with the exception of those originating from the RW and CF samples. Most of the total RNA is ribosomal in origin, and thus the profiles might simply reflect cell numbers. A rationale for

the RAP-PCR technique has been the assertion that ribosome per cell ratio is in most instances roughly proportional to the growth rate of bacteria (Wagner 1994). Furthermore, ~ 20% of the RNA may originate from mRNA (Aneja et al., 2004) and could provide some indication on gene expression. However, in our opinion there is no compelling evidence in the literature that this form for profiling is a good indicator of changes in total metabolic activity, particularly when a population is sequentially sampled as in the present work, or that the contribution of the mRNA fraction is sufficient to be distinguished from other RNAs in the analysis of bacterial communities. Water passes quickly from stage to stage in the treatment train, and thus the RNA content of the cells may remain relatively stable. The technique may be useful for comparative analyses of distinct, spatially separated samples, such as the raw water from two plants, but seems to offer little in studies of the present kind.

3.8 Sequence accession numbers

Sequences included in the study have the following GenBank Accession numbers: Fungal isolates: KR813020 - KR813025, bacterial isolate: KR813026, clones: KR813027 - KR813305.

4.0 Conclusion

The Vansjø water-treatment plant is a state-of-the art facility employing multiple purification and disinfection stages which together produce a drinking water of high hygienic quality. We hypothesized that individual stages would contribute in characteristic fashion to the final production of the drinking water, and that this might be reflected in the analyses performed. This was the case. The GAC filtration stage compromised water-quality with respect to microbial content. Furthermore, we show that especially post-chlorination (the ultimate treatment stage) and CF, the first treatment stage, exert a profound and positive effect on water quality and are of pivotal importance in achieving the plant's goals. The present study supports the contention that the drinking water is of pristine quality and extends the range of parameters supporting this contention beyond classical colony counts which reveal only a small percentage of the microbes present. We conclude that detailed and multi-faceted investigations of the microbiological changes at each step should be useful in trouble-shoooting, maintenance and improvement of treatment facilities. Few similar studies in the literature exist, and most are end-point investigations (ie, examining only the drinking water). As more data become available, particularly with regard to clarification employing flotation, it will be informative to see if similar trends to those reported here emerge.

640

641 **5.0 REFERENCES**

642

643 Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment
644 search tool. J. Mol. Biol. 215, 403-410.

645

646 Aneja, M.K., Sharma, S., Munch, J.C., Schloter, M., 2004. RNA fingerprinting: a new
647 method to screen for differences in plant litter degrading microbial communities. J.
648 Microbiol. Methods 59, 223-231.

649

650 Bahr, M., Hobbie, J.E., Sogin, M.L., 1996 Bacterial diversity in
651 an arctic lake: a freshwater SAR11 cluster. Aquat. Microb. Ecol. 11 (3), 271−277.

652

653 Berman, T., Kaplan, B., Chava, S., Viner, Y., Sherr, B.F., Sherr, E., 2001. Metabolically
654 active bacteria in Lake Kinneret. Aquat. Microb. Ecol. 23 (3), 213-224.

655

656 Boulos, L., Prevost, M., Barbeau, B., Coallier, J., Desjardins, R., 1999. LIVE/DEAD
657 BacLight: application of a new rapid staining method for direct enumeration of viable
658 and total bacteria in drinking water. J. Microbiol. Methods. 37 (1), 77-86.

659

660 Burkert, U., Warnecke, F., Babenzien, D., Zwirnmann, E., Pernthaler, J., 2003. Members of
661 readily enriched b-proteobacterial clade are common in surface waters of a humic lake.
662 Appl. Environ. Microbiol. 69, 6550–6559.

663

664 Chao, A., 1984. Nonparametric estimation of the number of classes in a population. Scand. J.

Statist 11, 265-270.

Charnock, C., Otterholt, E., Sundby, J., 2010. Investigations on changes in microbiological quality during production and distribution of drinking water at the Vansjø water-works. [Undersøkelse av forandringer i mikrobiologisk kvalitet under produksjon og distribusjon av drikkevann ved Vansjø interkommunale vannverk]. Vann 4, 485-495.

Ćirić, S.S., Petrović, O.V, Ilić, Z.Ž., Milosević, B.N, Spasić, Z.L., 2011. Examination of bacteriological status of surface fresh waters using direct and cultivation methods. Proc. Nat. Sci, Matica Srpska Novi Sad 121, 7-18.

Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., Tiedje, J.M., 2013. Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic acids Res. 42, (Database issue):D633-42. doi: 10.1093/nar/gkt1244

Demanou, J., Sharma, S., Njine, U., Weber, A., Wilke, B.M., Monkiedje, A., Munch, J.C., Schloter, M., 2006. Shifts in microbial community functions as a result of combined application of copper and mefenoxam. FEMS Microbiol. Lett. 260, 55–62.

Giovannoni, S.J., Tripp, H.J., Givan, S., Podar, M., Vergin, K.L., Baptista, D., Bibbs, L., Eads, J., Richardson, T.H., Noordewier, M., Rappe, M.S., Short, J.M., Carrington, J.C., Mathur, E.J., 2005. Genome streamlining in a cosmopolitan oceanic bacterium. Science 309, 1242–1245.

689

690 Godoy-Vitorino, F., Goldfarb, K.C., Brodie, E.L., Garcia-Amado, M.A., Michelangeli, F.,
691 Dominguez-Bello, M.G., 2010.Developmental microbial ecology of the crop of the
692 folivorous hoatzin. ISME J. 4, 611–620.

693

694 Good, I.J., 1953. The population frequencies of species and the estimation of population
695 parameters. Biometrika, 40 (3-4), 237-264.

696

697 Hammes, F., Berney, M., Wang, Y., Vital, M., Köster, O., Egli, T., 2008. Flow-cytometeric
698 total bacterial parameter for drinking water treatment processes. Water Res. 42 (1),
699 269-277.

700

701 Hijnen, W.A.M., Beerendonk, E.F., Gerriet Medema, G.J., 2006. Inactivation credit of UV
702 radiation for viruses, bacteria and protozoan (oo) cysts in water: a review. Water Res.
703 40 (1), 3-22.

704

705 Hill, T.C.J., Walsh, K.A., Harris, J.A., Moffett, B.F., 2003. Using ecological diversity
706 measures with bacterial communities. FEMS Microbiol. Ecol. 43 (1), 1-11.

707

708 Hoefel, D., Monis, P.T., Grooby, W.L., Andrews, S., Saint, C.P., 2005. Profiling bacterial
709 survival through a water treatment process and subsequent distribution system. J.
710 Appl. Microbiol. 99 (1), 175-186.

711 Juneja, D., Borah, A.K., Nasa, P., Singh, O., Javeri, Y., Dang, R. 2011. Candida sake
712 candidaemia in non-neutropenic critically ill patients: a case series. Crit. Care

713 Resusc.13(3), 187-191.
714
715 Kaarela, H., Härkki, O.E., Palmroth, M.R.T., Tukanen, T.A., 2015. Bacterial diversity and
716 active biomass in full-scale granular activated carbon filters operated at low water
717 temperatures. Environ. Technol. 36 (6), 681-692
718
719 Kahlisch, L., Henne, K., Gröbe, L., Brettar, I., Höfle, M.G. 2012. Assessing the viability of
720 bacterial species in drinking water by combined cellular and molecular analyses.
721 Microb. Ecol. 63 (2), 383-397.
722
723 Khan, Z., Mokkaddas, E., Ahmad, S., Burhamah, M.H., 2011. Isolation of Cryptococcus
724 magnus and Cryptococcus chernovii from nasal cavities of pediatric patients with
725 acute lymphoblastic leukemia. Med. Mycol. 49 (4), 439-443.
726
727 LeChevallier, M.W., Au, K.K., 2004. Water treatment and pathogen control: process
728 efficiency in achieving safe drinking water. IWA Publishing and World Health
729 Organization, London, UK.
730
731 McCammon, S.A., Bowman, J.P., 2000. Taxonomy of Antarctic Flavobacterium species:
732 description of Flavobacterium gillisiae sp. nov., Flavobacterium tegetincola sp. nov.,
733 and Flavobacterium xanthum sp. nov., nom. rev. and reclassification of
734 [Flavobacterium] salegens as Salegentibacter salegens gen. nov., comb. nov. Int. J.
735 Syst. Evol. Micr. 50 (3), 1055-1063.
736
737 Morris, R.M., Rappé, M.S., Connon, S.A., Vergin, K.L., Siebold, W.A., Carlson, C.A.,

738 Giovannoni, S.J., 2002. SAR11 clade dominates ocean surface bacterioplankton
739 communities. Nature, 420 (6917), 806-810.
740
741 Nielsen, J.L., Klausen, C., Nielsen, P.H., Burford, M., Jørgensen, N.O.G., 2006.
742 Detection of activity among uncultured Actinobacteria in a drinking water reservoir.
743 FEMS Microbiol. Ecol. 55, 432-438.
744
745 O'Connell, S.P., Garland, J.L. 2002. Dissimilar response of microbial communities in Biolog
746 GN and GN2 plates. Soil Biol. Biochem. 34 (3), 413-416.
747
748 Otterholt, E., Charnock, C., 2011a. Microbial quality and nutritional aspects of Norwegian
749 brand waters. Int. J. Food Microbiol. 144 (3), 455-463.
750
751 Otterholt, E., Charnock, C., 2011b. Identification and phylogeny of the small eukaryote
752 population of raw and drinking waters. Water Res. 45 (8), 2527-2538.
753
754 Pinto, A.J., Xi, C., Raskin, L., 2012. Bacterial community structure in the drinking water
755 microbiome is governed by fitration processes. Environ. Sci. Technol. 46 (16), 8851-
756 8859.
757
758 Reasoner, D.J., Geldreich, E.E., 1985. A new medium for the enumeration and
759 subculture of bacteria from potable water. Appl. Environ. Microbiol. 49, 1-7.
760
761
762 Shannon, C.E., 1948. A mathematical theory of communication. Bell System Technical
763 Journal 27, 379–423; 623-656.
764

765 Sharma, S., Aneja, M-K., Mayer, J., Schloter, M., Munch, J.C., 2004. RNA fingerprinting of
766 microbial community in the rhizosphere soil of grain legumes. FEMS Microbiol. Lett.
767 240, 181-186
768
769 Sharma, S., Szele, Z., Schilling, R., Munch, J.C., Schloter, M., 2006. Influence of freeze-thaw
770 stress on the structure and function of microbial communities and denitrifying
771 populations in soil. Appl. Environ. Microbiol. 72 (3), 2148-2154.
772
773 Staley, J.T, Konopka, A., 1985. Measurement of *in situ* activities of nonphotosynthetic
774 microorganisms in aquatic and terrestrial habitats. Annu. Rev. Microbiol 39, 321-346.
775
776 Stefanowicz, A., 2006. The biolog plates technique as a tool in ecological studies of
777 microbial communities. Polish J. Environ. Stud. 15 (5), 669-676.
778
779 Wagner, R., 1994 The regulation of ribosomal RNA synthesis and bacterial cell
780 growth. Arch. Microbiol. 161, 100-106.
781
782 Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R., 2007. Naive Bayesian classifier for rapid
783 assignment of rRNA sequences into the new bacterial taxonomy. Appl. Environ,
784 Microbiol. 73, 5261-5267.
785
786
787
788
789

Table 1 – colony counts, protozoa and GN2 metabolic fingerprints at each treatment stage

Sample	Sabouraud agar (CFU in 250mL)	Protozoa: Presence(P) Absence (A)	R2A agar (CFU mL^{-1})	Blood agar (CFU mL^{-1})	GN2 positive tests
RW	250/110	P/P (flagellates, ciliates, amoebae)	1670/1735 (920)*	ND/340 (50)	91/91
CF	20/40	A/A	2600/700 (1400)	ND/10 (10)	64/42
2MF	5575/63	A/A	1710/1270 (1000)	ND/<10 (<10)	65/12
BUV	425/25	A/A	1510/1260 (615)	ND/20 (<10)	68/20
AUV	98/2	A/A	45/40 (75)	ND/<10 (<10)	32/3
GAC	1/0	P/A	205/75 (310)	ND/<10 (10)	15/9
DW	0/0	A/A	<10 /<10 (<10)	ND/<10 (<10)	0/0

Results from R1 of testing are given before '/' and R2 after. ie, R1/R2.

< 10 values = None detected in 0.1mL

* equivalent data from Charnock et al., (2010) is given in parentheses. NB, CFU mL^{-1} on R2A, was counted after 3 weeks.

Table 2 – Richness, diversity and coverage data for clone library OTUs

Treatment stage	No. of clones	Number of OTUs		Chao1[a]		Good's index[b] (%)		Berger-parker index (d) = n_{max}/N		Margalef index (D_{Mg}) = (S-1)/lnN		Shannon-Wiener index (H') = $\Sigma p_i ln p_i$ [c]	
	Cut-off (%)	97	98	97	98	97	98	97	98	97	98	97	98
RW	39	27	27	70	70	49	49	0.10	0.10	7.1	7.1	3.3	3.3
CF	45	20	22	125	142	67	64	0.33	0.24	5.0	5.5	2.5	2.7
2MF	27	10	14	31	37	74	63	0.33	0.33	2.7	3.9	1.9	2.3
BUV	30	16	17	38	63	60	53	0.37	0.37	4.4	4.7	2.3	2.4
AUV	44	25	26	40	43	66	64	0.11	0.11	6.3	6.6	3.0	3.1
GAC	47	28	30	57	69	60	53	0.15	0.15	7.0	7.5	3.1	3.2
DW	47	25	25	83	83	57	57	0.34	0.34	6.2	6.2	2.7	2.7
Total	279	89	100	414	408	76	73	0.20	0.20	16	18	3.4	3.7

n_{max} = number of clones in most abundant OTU. N = total number of clones in sample

S = number of OTUs in the sample

[a]Chao1 = $S_{Chao1} = S_{obs} + n_1 (n_1-1)/2(n_2+1)$. Where S_{obs} is the observed number of species, and n_1 and n_2 are the number of OTUs with only one sequence and 2 sequences respectively (Chao, 1984).

[b] Coverage: sum of probabilities of observed classes calculated as (1-(n/N))×100 where n is the number of singleton sequences and N is the total number of sequences (Good, 1953).

[c]p_i =proportion of clones in the ith OTU (estimated using n_i/N) (Shannon, 1948).

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828