Molecular diversity of drinking water microbial communities: a phylogenetic approach

Randy P. Revetta,* Jorge W. Santo Domingo, Catherine Kelty,
U. S. Environmental Protection Agency, Office of Research and Development,
National Risk Management Research Laboratory, Cincinnati, OH 45268

Ben Humrighouse, Daniel Oerther, Regina Lamendella, Department of Civil and Environmental Engineering, University of Cincinnati, Cincinnati, OH 45221-0071

Minna Keinanen-Toivola, Prizztech Ltd., DWI Finland, Rauma, Finland

Margaret Williams, Centers for Disease Control and Prevention, Atlanta, GA

* Corresponding author: revetta.randy@epa.gov

ABSTRACT

The microbiological quality of drinking water is assessed using culture-based methods that are highly selective and that tend to underestimate the densities and diversity of microbial populations inhabiting distribution systems. In order to better understand the effect of different disinfection treatments on the bacteria in water distribution systems (WDS) we have used 16S rDNA PCR-based techniques. 16S rDNA clone libraries were developed using DNA extracted from samples collected at different times and locations in a metropolitan distribution system derived from different source waters. Since it is possible for DNA to persist in the environment after cell death we also used rRNA (instead of DNA) as the target molecule to study the active bacterial fraction in drinking water biofilms. Phylogenetic analyses and sequence comparisons with existing databases revealed that α- and β-Proteobacteria were among the most predominant bacterial groups identified in both biofilm and planktonic communities. Clones closely related to *Mycobacterium* spp. and *Legionella* spp. were also obtained in these studies. While members of the *Mycobacterium* and *Legionella* genera are known to be pathogenic, the public health relevance of these clones has yet to be determined. Differences in biofilm community structure between disinfection treatments were evident as biofilms exposed to chloramine primarily consisted of *Mycobacterium* spp. and *Dechloromonas* spp. In contrast, a variety of α- and β-Proteobacteria dominated the DNA-based clone libraries of biofilm receiving no disinfectant and biofilm exposed to chlorine. Sequence analysis of RNA-based clones derived from biofilms exposed to chlorine suggested that the active bacterial fraction consisted of a few dominant bacterial groups related to *Nevskia ramosa*, a member of the γ-Proteobacteria group frequently found in the air-water interphase (*i.e.*, neuston). Some of the bacterial groups identified in the latter clone libraries are affiliated with yet to be cultured organisms suggesting that a significant fraction of drinking water communities cannot be studied using culturing methods. The results from these studies suggest that we have a limited understanding of the molecular diversity and population dynamics of WDS microbial communities. Future studies will focus on expanding sequencing databases to more accurately characterize potential differences between disinfection treatments and to better understand the microbial diversity in drinking water systems.

KEYWORDS

16S rDNA, biofilms, clone libraries, phylogenetic analysis

INTRODUCTION

Drinking water in developed countries is typically of excellent initial quality although changes in the bacterial composition can occur during its residence in the water distribution system (WDS). For example, intrusion due to loss of pipe structural integrity and shedding of bacteria from biofilms found on pipe surfaces in the distribution system can introduce changes in the drinking water microbiota. More importantly, WDS biofilms have been shown to mitigate the effectiveness of commonly used disinfectants and as a result increase the survival of pathogens and their growth potential in distribution systems (LeChavallier et al. 1996; Williams & Braun-Howland, 2003.

The microbiological quality of drinking water is assessed using culture-based methods that are highly selective and that tend to underestimate the densities and diversity of microbial populations inhabiting distribution systems. One explanation for this phenomenon relates to the fact that it very difficult to simulate in vitro the in vivo physicochemical interactions that are part of natural microbial assemblages. Additionally, stressful environmental habitats like drinking water contain many injured bacteria that cannot grow well on artificial media. As a result, recent descriptions of WDS microbial communities have incorporated nucleic acid-based approaches that avoid the requirement of culturing microbial populations prior to taxonomic identification. These culture-independent methods have included internal transcribed spacer (ITS) fingerprints (Pozos et al., 2004), terminal restriction fragment length polymorphisms (T-RFLP) (Martiny et al., 2003), 16S rRNA gene surveys (Santo Domingo et al., 2003; Williams et al., 2004; Keinanen-Toivola et al., 2006) and metagenomic surveys (Schmeisser et al., 2003). Our laboratory has focused on the sequencing analysis of ribosomal RNA (rRNA) genes (i.e., rDNA) to study natural microbial communities as this approach can be used for the identification of microbial populations. Since 16S rDNA sequences can discriminate between closely related species it also provides the opportunity to more accurately estimate in vivo microbial diversity.

Herein we report on the application of a 16S rDNA-based clone library approach to examine the planktonic microbial community structure in a distribution system simulator receiving different disinfection treatments and in a metropolitan WDS receiving different water sources. In addition, we also report on the microbial communities inhabiting WDS biofilms exposed to different disinfection treatments and on the identification of active bacterial communities in a model drinking water biofilm system. The results suggest that drinking water microbial communities are indeed complex and that a significant fraction has yet to be identified using conventional culture-based approaches. The need for incorporating molecular surveys when studying the microbial dynamics of drinking water and its relevance to public health and disinfection treatment is discussed.

METHODOLOGY

Distribution System Simulator

Water samples were obtained from the Distribution System Simulator (DSS) located at the EPA Test and Evaluation (T&E) facility in Cincinnati, OH. The DSS consists of six independent ductile iron pipe loops supplied with tap water from the Greater Cincinnati Water Works. The pipe loops have been in operation for over 7 years, with exposure to various chlorine and monochloramine residuals. Additional details regarding the DSS, including data on physicochemical parameters, are described by Meckes et al. (1999). Samples from the DSS were collected from one of the loops receiving Cincinnati distribution water containing a 1 ppm free-chlorine residual amended with NH_3 to create a monochloramine (NH_2Cl) residual before entering the loop system for a 24-h residence time. Feed and discharge samples were filtered onto polycarbonate filters, and further processed as described below.

WDS Point of Use Sampling

Water samples used in our studies were collected from 8 sites at cold water taps (point of use) in private, public or commercial buildings throughout a metropolitan drinking water distribution system. The service area is supplied by two treatment facilities with different source water and different treatment processes. The source for WDS1 is groundwater while the source of WDS2 is surface water. Samples were obtained directly from faucet heads after they were run with the cold valve completely open for 5 minutes. This was performed to flush out idle water from the pipes within the structure and thus get representative samples from each sampling point. A total of 5 liters were collected in sterile polypropylene (Nalgene) bottles from each sampling location and transported to the laboratory in coolers. Samples were filtered onto polycarbonate filters, and stored at -80°C within three hours of collection. Biomass on the filters was used for DNA extractions.

Biofilm Studies

Biofilms were grown in annular reactors (BioSurface Technologies Corp., Bozeman, Montana) each connected to a 20 L carboy reservoir containing city of Cincinnati distribution water obtained from the laboratory cold tap. The reactors consisted of inner and outer cylinders with a wetted surface area of $0.19 m^2$ and a liquid volume of 0.6 L. Biofilms were developed on polycarbonate slides at room temperature for 10 weeks following the operator's manual. The inner cylinder of the reactor rotated at a rate to mimic a water velocity of 0.23 m s^{-1}. At each sampling timepoint, polycarbonate slides were removed aseptically from reactors and transferred to sterile glass tubes containing 40mL of EmbryoMax Ultrapure Water (Specialty Media, Phillipsburg, NJ) and 5 g of sterile glass beads (diameter 3 mm, PGC Scientific, Frederick, MD). The biofilms were removed from the slide surface by vortexing for 5 min at maximum speed (Vortex Genie2, Bohemia, NY). Alternately, slides were placed in sterile 50ml conical tubes containing 30ml sterile phosphate buffered saline, and biofilm was aseptically scraped from the slide using a sterile Teflon policeman, followed by vortexing. Supernatants containing the

biofilm were then filtered onto 0.22 μm pore size polycarbonate filters (Osmonics Inc., Minnetonka, MN). Membranes were stored at -80°C until used in molecular studies.

Culturing and Molecular Techniques

The artificial medium R2A was used to isolate bacteria from drinking water. This medium has been shown to isolate a higher number of bacteria from low nutrient (oligotrophic) environments (Reasoner and Geldreich, 1985). Samples were collected using sterile containers and water aliquots were spotted directly onto R2A agar plates and processed as described elsewhere (Santo Domingo et al, 2003). Colonies were randomly selected and resuspended in sterile water. Aliquots (1 μl) from resuspended cells were used as DNA template for the amplification of 16S rDNA using the polymerase chain reaction (PCR). PCR products were then sequenced as described below.

DNA extraction from polycarbonate filters was performed using an UltraClean™ Soil DNA kit following the manufacturer's instructions (MoBio Laboratories, Solano Beach, CA). DNA extracts were used to amplify a partial region of the community 16S rDNA which was then used to generate clone libraries. RNA was extracted with TRIZOL® Reagent following the manufacturer's instructions (Invitrogen, Carlsbad, CA). Extracted RNA was treated enzymatically with DNase I to remove coextracted DNA using DNA-free™ and TURBO DNA-free™ kits (Ambion, Austin, TX). A two-step RT-PCR procedure was used to generate products for RNA-based clone libraries. First, DNase I-treated RNA was used as a template in RT-step experiments using an EndoFree RT™ kit (Ambion) and 16S rRNA gene eubacterial primer 685R. Transcript (i.e. cDNA) generated in the RT step was then used as a template in a PCR step using 16S rRNA gene eubacterial primers 27F and 685R (Lane, 1991). DNA extracts from filters or resuspended cells were used in PCR assays using 27F and 685R, 787R, 1525R primers.

Amplification products were purified with QIAquick PCR Purification kits (Qiagen, Valencia, CA) and cloned using TOPO TA Cloning kits (Invitrogen). Clones were sequenced in both directions on a BaseStation 51 DNA Fragment Analyzer (MJ Research, Waltham, MA), or on an ABI 3730XL DNA analyzer (Applied Biosystems, Foster City, CA). Sequencher 4.6 (Gene Codes Corp., Ann Arbor, MI) was used to edit the sequences, and BLAST (Autshul et al., 1997) was used to identify closest matching sequences in the NCBI databases. Sequences were examined for chimeras with Chimera Check (Cole et al., 2003). Sequence alignment and phylogenetic analyses were performed with ARB (Ludwig et al. 2004) (http://www.arb.de.vu) using the neighbor-joining algorithm of Saitou and Nei (1987) with Felsenstein correction.

Taxonomic Diversity

Diversity within the overall planktonic and biofilm libraries was analyzed by rarefaction using the analytical approximation algorithm of Hurlbert (1971), and the software program aRareFactWin by Holland (1998). Coverage values were calculated by the equation $C = 1 - (n/N) \times 100$, where n is the number of operational taxonomic units (OTU), N is the number of clones examined, and C is the percent coverage. OTUs were defined as clones that shared 97% or greater sequence similarity. Sequences included in this paper and deposited in the NCBI database are listed in Table 1.

Table 1 – Summary of Sequences Analyzed

Sample Description	Sample Prefix	Collection Date	No. of sequences	NCBI Accension number
Cincinnati Tap Water (HOCl)	HOClCi	2002	74	AY328550 – AY328623
DSS feed water (NH_2Cl)	DSSF	2002	77	AY328624 – AY328700
DSS discharge water (NH_2Cl)	DSSD	2002	98	AY328701 – AY328798
WDS-2 surface water	WDS-2	2005/2006	328	This study
WDS-1 Ground water	WDS-1	2005/2006	298	This study
DSS R2A isolates	OR	2002	14	AY328799 – AY328812
	Y	2002	15	AY328813 – AY328827
	MA	2002	10	AY328828 – AY328837
	MB	2002	14	AY328838 – AY328851
RNA Biofilm (HOCl)	ARRNA	2005	52	DQ149874 – DQ149897
DNA Biofilm (HOCl)	ARDNA	2005	32	DQ149898 – DQ149917
DNA Biofilm (HOCl)	ARmtp	2005	109	This study
B1N Biofilm (No disinfectant)	B1N	2003/2004	9	AY957886 – AY957894
B2C Biofilm (HOCl)	B2C	2003/2004	10	AY957895 – AY957904
B3N Biofilm (No disinfectant)	B3N	2003/2004	36	AY957905 – AY957941
B4M Biofilm (NH_2Cl)	B4M	2003/2004	9	AY957942 – AY957950

RESULTS

To further understand and assess microbial diversity in drinking water, clone libraries were constructed using DNA and RNA from model biofilms exposed to different disinfection treatments. In addition, clone libraries were constructed using planktonic samples from a distribution system simulator exposed to different disinfection treatments, and from sites within a WDS receiving groundwater and surface water as the primary sources of water. Overall, phylogenetic relationships were inferred from 1185 partial 16S rRNA gene sequences (620 positions) of clones derived from model biofilms (22%), the DSS (25%), and from sites (53%) throughout a metropolitan WDS (Figure 1).

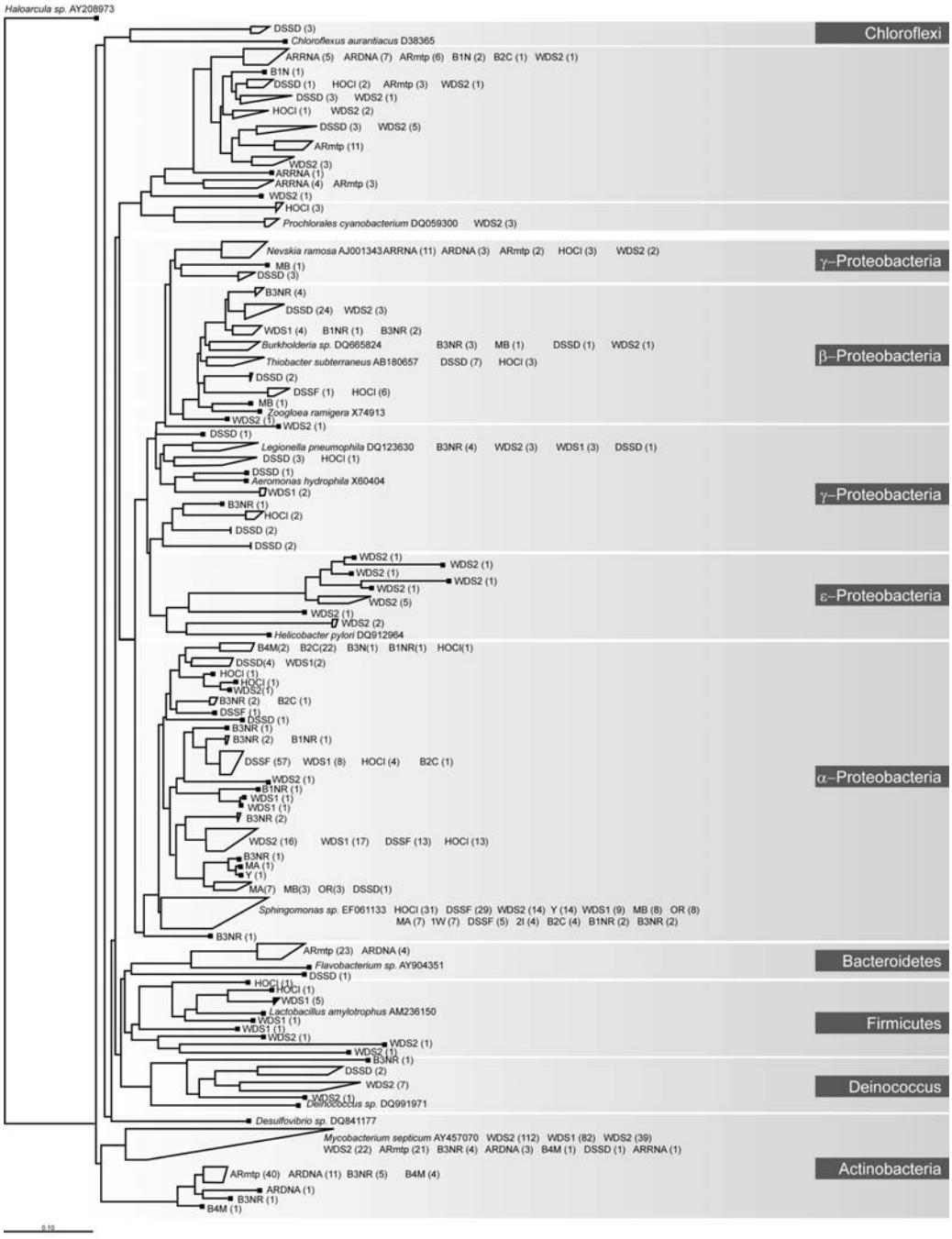

Figure 1 – Schematic representation of the distribution of 16S rRNA gene sequences derived from drinking water biofilm and planktonic samples. The tree was constructed with a neighbor-joining algorithm using Felsenstein correction. Information in parentheses refers to the number of sequences per library type (see Table 1) as compared to cultured reference sequences. The 16S rRNA gene sequence for *Haloarcula sp.* was used as the outgroup. HOCl = tap water; DSSF = distribution system simulator feed water; DSSD = distribution system simulator discharge water; WDS2 = distribution system - surface water source; WDS1 = distribution system - ground water source; OR/Y/MA/MB = distribution system simulator R2A isolates; ARRNA = model biofilm RNA; ARDNA = model biofilm DNA; ARmtp = model biofilm DNA; B1N = model biofilm no disinfectant; B2C = model biofilm chlorinated; B3N = model biofilm no disinfectant; B4M = model biofilm monochloramine

Phylogenetic Diversity of Drinking Water Bacteria in a Distribution System Simulator

Overall, 302 sequences were obtained from planktonic samples collected at the EPA DSS. Chlorinated distribution system water (HOClCi) contained 51% α-Proteobacteria closely related to *Sphingomonas*, *Brevundimonas*, *Afipia*, *Blastomonas*, *Hyphomicrobium*, *Methylocystis*, *Bradyrhizobium* species, and an endosymbiont of *Acanthamoeba* that is closely related to *Caedibacter caryophilus*. These samples also containded clones closely related to cyanobacteria, firmicutes, β- (*Burkholderia* sp.) and γ-Proteobacteria (*Nevskia* sp.). The chloraminated DSS feed water (DSSF) contained some of the same α-Proteobacteria genera found in HOClCi, however, the majority of clones were closely related to *Hyphomicrobium* sp. In contrast, 16S rDNA clones from the DSS discharge water (DSSD) showed a greater diversity of cyanobacteria, and α-, β-, γ-Proteobacteria than the DSSF water. Other bacterial genera noted in this clone library were *Nitrospira*, *Actinobacteria*,and *Planctomycetacia*.

Most of the bacterial strains isolated (i.e., 94%) from DSSD water were α-Proteobacteria, the majority belonging to the *Brevundimonas*, *Caulobacter*, *Sphingomonas*, or *Novosphingobium* sp. A few isolates were identified as *Porphyrobacter*, *Afipia*, or *Phaeospirillum* sp. Three isolates were closely related to *Propionivibrio limicola* and *Limnobacter thiooxidans* of the β-Proteobacteria.

Microbial Characterization of Drinking Water Systems Receiving Groundwater and Surface Water

Overall, 626 sequences were isolated from 8 sites at cold water taps (point of use) in private, public or commercial buildings within a metropolitan water distribution system. Four sample sites were in the service area supplied by a treatment facility receiving groundwater, and 4 sites in the service area supplied by a treatment facility receiving surface water. Overall, *Mycobacterium sp.* and α-Proteobacteria represented nearly 64% and 28% of the total clones examined, respectively. Other bacterial groups identified included members of the genera *Legionella*, *Pseudomonas*, and *Agrobacterium*. Approximately 10% of the sequences analyzed in this study showed less than 97 % identity with sequences in the currently available databases. In addition, nearly a third of sequences were closely related to uncultured bacteria. A significant number of sequences obtained in this study were related to the following mycobacterial species: *M. mucogenicum*, *M. massiliense*, *M. sacrum*, *M. gordonae*, and *M. gadium* (with sequence similarities of 98% and higher). Moreover, drinking water samples originating from the distribution system receiving groundwater as the water source contained approximately three times the number of *M. sacrum*- and *M. massiliense*-like sequences than the samples obtained from the system receiving surface water. Most of the α-Proteobacterial sequences were closely related to *Sphingomonas, Porphyrobacter, Hyphomicrobium,* and *Methylocystis* sp. In addition, some of the clones were closely related to sequences from yet to be cultured rhizosphere bacteria and to an endosymbiont of *Acanthamoeba*. Interestingly, while *Sphingomonodaceae* sequences were obtained from both distribution systems, the distribution system using groundwater as the main water source contained 5 times more sequences than the system using surface water. *Methylocystis* and *Porphyrobacter* species are not normally associated with drinking water and perhaps are transitory members of this habitat. The β-proteobacteria sequences obtained are

closely related to *Delftia tsuruhatensis* (100% sequence similarity), a bacterium isolated from activated sludge, and *Aquaspirillum delicatum* (97% sequence similarity) (Shigematsu et al., 2003).

Population Diversity in Model Potable Water Biofilms Receiving Chlorine or Chloramine Residual

Overall, 64 clones were analyzed over a 10-week incubation period; 70% from control biofilms (no disinfectant), 14% from biofilm grown under exposure to chloramine, and 8% from biofilm under exposure to chloramine. The biofilm grown under exposure to chloramine residuals was dominated by organisms closely related to *Dechloromonas* species and *Mycobacterium cosmeticum*. Only one clone sequence from the chloramine biofilm was not affiliated with the two latter genera but with *Bradyrhizobium japonicum*, an α-Proteobacterium instead. In contrast, the other biofilms contained a more diverse microbial community that included several clones that clustered with other α Proteobacteria in addition to *Bradyrhizobium* sp. (i.e. *Sphingomonas*, and *Hyphomicrobium*). While the control 1 biofilm and chlorine biofilm contained sequences mostly from α- and β-Proteobacteria, the control 2 biofilm contained the most diversity in its cloned sequences, including sequences related to α-, β-, γ-Proteobacteria, and Actinobacteria. In addition, clone B3NR69D25 from the control 2 biofilm (day 69) showed 97% similarity with *Caedibacter caryophilus*, a macronuclear endosymbiont of *Paramecium caudatum*, while clone B3NR69D10 obtained from the same sample showed 95% similarity with a sequence belonging to a symbiont of *Amoeba proteus* (recently classified as *Legionella jeonii*; Park et al. 2004) and 94% similarity to *L. pneumophila* str. Paris (CR628336). No 16S rDNA sequences similar to those of *Legionella* sp. were detected in the control 1 biofilm reactor or in the biofilms exposed to chlorine or chloramine. The relative species richness values indicate that OTU coverage of the sample diversity was similar for the control 1 and chlorine biofilms (0.099 and 0.11 respectively). The highest relative species richness was found in the control 2 biofilm, with similar values for the samples obtained on days 56 and 69 (0.23 and 0.19). These values were approximately twice as high as the control 1 biofilm species richness. By contrast, the chloraminated biofilm samples contained the lowest species richness on both days (0.042 and 0.026). The development of culturable heterotrophic populations in biofilms receiving residual disinfectant was slowed, but not stopped under exposure to chlorine or chloramine, as measured by R2A plate counts. Levels of culturable heterotrophs in biofilms receiving residual disinfectant slightly exceeded those in the biofilms receiving no residual within the first month of incubation. The highest CFU levels were observed in biofilms from experiment 2 after 10 weeks' incubation, at approximately 5×10^6 and 7×10^6 CFU cm^{-2} in the control 2 and chloramine biofilms, respectively. In the reactor which received chlorinated water, the density of heterotrophic bacteria stabilized at 10^5 CFU cm^{-2}. However, these populations declined between week 4 and week 10 in the control 1 biofilm. Amoebae were isolated from the non-disinfected control 1 and 2 biofilms after 1 month, while no amoebae were obtained from the biofilms receiving chlorine or chloramine residual. The density of heterotrophic bacteria did not decline in the control 2 biofilm, even though amoebae were present in the system.

Identification of Active Bacterial Communities in a Model Drinking Water Biofilm System Using 16S rRNA-based Clone Libraries

Overall, 193 sequences derived from 16S rDNA-based and rRNA-based clone libraries were analyzed in biofilm grown under exposure to chlorine disinfectant over a 20-week incubation period. Thirty percent of the sequences analyzed were derived from rRNA-based clone libraries. More than half (53.8%) of the identified sequences in the rRNA-based clone library were closely related to *Nevskia ramosa*, a γ-Proteobacteria. In addition, nearly 30% of clones were closely related (i.e. > 97% sequence identity) to sequences of uncultured bacteria recovered from a recent drinking water study that used DNA as the target molecule to develop 16S rRNA gene libraries (Williams et al., 2004). Several clones (15.4%) were not closely related to any sequences present in publicly available databases while one 16S rRNA-based clone was classified as *Mycobacterium* spp. based on sequence similarity. An rDNA-based clone library was also developed to compare the phylogenetic affiliation of active bacteria vs. clones recovered when using DNA as a template. Analogous to the rRNA-based library, members of the genus *N. ramosa* (32.3%) and *Mycobacterium* spp. (12.9%) were represented in the rDNA-based library. In contrast to the rRNA-based library, *Dechloromonas* spp. was also identified as part of the rDNA-based clone library. Phylogenetic analyses indicated that approximately 25% of the clone sequences were closely related to uncultured bacteria and to clones previously recovered from Cincinnati drinking water biofilm samples (Williams et al., 2004, 2006). Other clones were closely related to sequences recovered from other environmental samples (19.4%), whereas a small group of clones (i.e. 3.2%) did not have close representation (i.e. < 94% similarity) in the publicly available databases.

DISCUSSION

Phylogenetic analyses using16S rDNA-based and rRNA-based cloned sequences were used to study model drinking water biofilm systems exposed to different disinfectant treatments. While the presence of disinfectant slowed model biofilm formation initially as measured by culturable heterotrophic populations (data not shown), it did not completely inhibit biofilm formation as bacterial levels were restored to original levels after 4 weeks of reactor operation and regardless of the disinfectant treatment. Phylogenetic analyses however showed that the microbial community structure of control biofilms and biofilms exposed to disinfectants was different. Specifically, 10-week old biofilms developed under chloramine exposure were dominated by *Dechloromonas*-like bacteria and *M. cosmeticum* (Cooksey et al. 2004). The occurrence of *Dechloromonas* and *Mycobacterium* species in biofilms and water samples receiving the same water source suggests they are a normal component of this drinking water system, and demonstrates their ability to survive conventional disinfection treatment.

Interestingly, *Dechloromonas* were not observed in the RNA-based clone libraries, suggesting that they are not among the most active drinking water bacterial populations. Moreover, it is possible that the *Dechloromonas* clones were recovered from dead cells, in turn indicating that their relevance to biofilm formation is not clear. The RNA-based clone libraries were dominated (54%) by sequences closely related to *Nevskia ramosa*, a γ-Proteobacterium. Assuming that the presence of RNA is indicative of potential cellular metabolic activity, these results suggest that *Nevskia*-like bacteria are not only capable of surviving, but perhaps even growing under the

challenging environmental conditions present in drinking water This is perhaps not surprising as they have been frequently found in harsh environments like the air-water interphase (Babenzien, 1989) and in several aquatic habitats. Samples from a wide variety of aquatic environments (ponds, lakes, rivers, ditches) were analyzed for these rosette-forming neustonic bacteria. These bacteria form hydrophobic surface films, but only in the absence of combined nitrogen (Styrmeyer et al., 1998; Pladdies et al., 2004). Sequences closely related (i.e. > 97% sequence identity) to uncultured bacteria constituted a significant portion (45%) of the rRNA-based clones libraries, further suggesting our limited understanding of the microbial diversity in drinking water systems.

The fact that most of the *Nevskia*, *Mycobacteria*, and uncultured bacterial clones were recovered from independent clone libraries generated using different 16S rRNA gene primers, and using both RNA and DNA as a target further suggests the ubiquity and potential active status of these bacterial groups in the studied water distribution system. Although their ecological role has yet to be determined, the abundance of uncultured bacteria suggests that they could be important members of drinking water biofilm communities. Our results also underline the importance of developing molecular-based methods that both detect and enumerate active uncultured bacteria in drinking water systems. Such methods will prove to be useful in studying the ecology and survival characteristics of these bacteria, which in turn will help us better understand potential risks associated with the common bacterial groups found in potable water.

Analyses of 16S rDNA cloned sequences were also compared between chlorinated water from a metropolitan water distribution system, chloraminated feed and discharge water of a DSS, and culturable heterotroph populations from the DSS discharge. Members of the α-Proteobacteria group were the predominant phylogenetic group observed following exposure to either free chlorine or monochloramines. In addition, most bacterial strains isolated from R2A plates belonged to the α-Proteobacteria, indicating the viability of these organisms in chloraminated distribution system water. Conversely, the fact that other proteobacterial groups and Gram-positive bacteria were observed in the clone libraries indicates the strong bias associated with culture-based techniques. While *Hyphomicrobium* sp. dominated in the chloraminated water, they were also present in the chlorinated water. Considering that this bacterial group has been isolated in many oligotrophic aquatic environments, it is reasonable to speculate they may play an important role in carbon utilization as this group is known to grow under carbon limited conditions (Corpe and Jensen 1996). A previous study of bacterial populations in the DSS demonstrated the predominance of α- and β-Proteobacteria within the culturable isolates and total community DNA (Santo Domingo et al. 2003). The highest percentage of β-Proteobacteria occurred in water with the lowest disinfectant residual, suggesting that these organisms are more sensitive than the observed α-Proteobacteria to relatively low levels of chlorine disinfectant exposure.

Phylogenetic analysis of clone sequences from various locations in a WDS illustrates many similarities with the studies conducted at the DSS. For example 16S rDNA sequences affiliated to mycobacteria, α-Proteobacteria, and *Legionella*-like organisms have been previously shown to be present in both drinking water biofilms and drinking water planktonic populations. Including this study, nearly identical sequences have been recovered in several independent studies from the same distribution system suggesting that these bacterial groups can be considered to be part

of the normal microbiota of this drinking water system. Interestingly, samples originating from the distribution system receiving groundwater as the water source contained approximately three times the number of *M. sacrum*- and *M. massiliense*-like sequences than the samples obtained from the system receiving surface water. This data suggests that some conditions in the distribution system might selectively favor the proliferation of some types of mycobacterial species. For example, previous studies have shown that the density of mycobacteria increases with the increase of AOC (Falkinham et al., 2001). Other factors that may positively influence mycobacterial densities are heterotrophic bacterial counts, iron, water retention times and turbidity (Falkinham et al., 2001).

Phylogenetic analysis of bacterial sequences indicated some interesting observations among the distribution of microbial groups in planktonic and biofilm environments. The sequences derived from planktonic samples were dominated by α-Proteobacteria (40%) and Actinobacteria (38%), and included a diverse group including β- and γ-Proteobacteria, Cyanobacteria, Firmicutes, and Nitrospirae (Table 2). In contrast, sequences derived from biofilm samples were less diverse and were dominated (49%) by sequences with no close matches or closely related to uncultured organisms only (Table 3).

Table 2 – Summary of planktonic bacterial sequence identification according to the closest matches to sequences in the NCBI databases.

	Tap Water (HOCl)	DSS feed water (NH_2Cl)	DSS discharge (NH_2Cl)	DSS R2A isolates	WDS-1 ground water	WDS-2 surface water	WDS-1 (1W)	WDS-1 (2I)	WDS-2 (3K)	WDS-2 (4B)
No. sequences analyzed	**74**	**77**	**98**	**53**	**146**	**165**	**88**	**64**	**90**	**73**
Unknowns *	37.7%		26.5%		8.9%	18.8%		15.6%	4.4%	1.4%
α-proteobacteria	51.4%	98.7	33.7	94.3	33.6%	24.9%	44.3%	14.1%	38.9%	1.4%
β-proteobacteria		1.3%	28.6%	5.7%		3.0%		6.2%		
γ-proteobacteria	4.1%		6.1%		7.5%	2.4%		3.1%		
Actinobacteria			2.0%		50%	46.7%	55.7%	51.6%	56.7%	95.8%
Cyanobacteria	4.1%					4.2%				
Firmicutes	2.7%							9.4%		1.4%
Nitrospirae			3.1%							

*** Sequences with no close matches (<97%), or closely related to uncultured organisms only.**

Table 3 – Summary of biofilm bacterial sequence identification according to the closest matches to sequences in the NCBI databases.

	AR-RNA (HOCl)	AR-DNA (HOCl)	ARmtp – DNA (HOCl)	B1N (No disinfectant)	B2C (HOCl)	B3N (No disinfectant)	B4M (NH_2Cl)
No. sequences analyzed	**52**	**32**	**109**	**9**	**10**	**36**	**9**
Unknowns *	44.2%	47.6%	56.0%	66.7%	50%	36.1%	22.2%
α-proteobacteria				22.2%	50%	27.8%	
β-proteobacteria		7.2%	24.7%	11.1%		25%	55.6%
γ-proteobacteria	53.9%	32.3%				2.8%	
Actinobacteria	1.9%	12.9%	19.3%			8.3%	22.2%

*** Sequences with no close matches (< 97%), or closely related to uncultured organisms only.**

Various statistical approaches have been developed to estimate microbial diversity, including diversity indices, estimated species richness, and sample diversity comparisons with rarefaction curves (Hughes et al., 2001). Rarefaction compares observed diversity (i.e. richness or number of types) among different sites, treatments or habitats that have been unequally sampled. Diversity of the biofilm and planktonic clone libraries were investigated by rarefaction analysis to evaluate whether the number of clones analyzed was sufficient to estimate the diversity in our samples. As illustrated in Figure 2, the calculated rarefaction curves did not reach saturation, indicating that analysis of an increasing number of clones would have revealed more OTU's in each sample. However, with the advent of high throughput sequencing technology it is reasonable to expect that several thousand additional clones will be analyzed, revealing additional insights into relative species abundances both between and among clones derived from planktonic and biofilm drinking water samples.

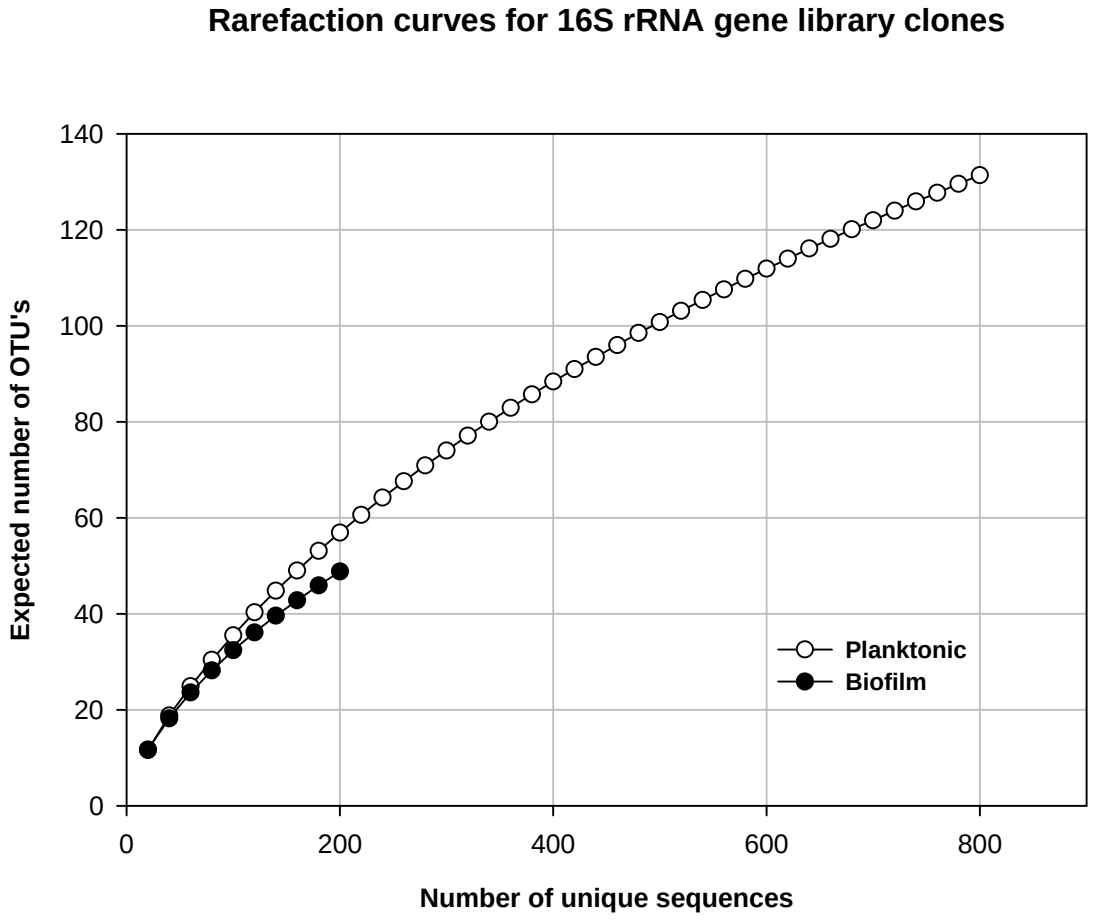

Figure 2 – Rarefaction curves for unique 16S rRNA gene library clones. Rarefaction curves were calculated using the analytical approximation algorithm described by Hurlbert (1971). Sequences were grouped into operational taxonomic units (OTU's) based on 97% or greater sequence similarity. The expected number of OTU's is plotted vs. the number of unique 16S rRNA clones. Rarefaction curves were calculated for planktonic (open circle), and biofilm (closed circle) clone libraries.

CONCLUSIONS

The sequence data associated with these studies supports the notion that microbial community structure is dynamic and that it can be impacted by differences in the water biogeochemistry, disinfection treatment, and type of water source. Indeed, it is clear that disinfection strategy impacts microbial community diversity and that certain bacterial groups tend to be relatively common members of WDS even if chlorine residual levels are maintained. Moreover, recent data generated in our laboratory and others (Keinänen-Toivola et al., 2006; Keinänen et al., 2003; Kalmbach et al., 1999; Williams et al. 2004) have clearly demonstrated that some mycobacteria populations in drinking water are potentially active or at least capable of growing. It should be noted that the recovery of certain mycobacteria and *Legionella* 16S rDNA sequences from drinking water DNA extracts does not mean that these specific bacterial populations can cause a serious threat to the consumer. However, these data suggest that it may be possible for some of the pathogenic close relatives to withstand disinfection treatment technologies, particularly within the distribution system. In addition, these results suggest that there may be some differences of the drinking water treatment unit processes in the area of study, pertaining to the removal of these organisms. These findings may warrant a closer investigation of mycobacteria behavior pertaining to drinking water distribution systems. Future studies should focus on the specific detection and quantification of active mycobacteria, specifically those of potential public health risk value. In addition, metagenomic approaches have been used to assess phylogenetic diversity (Schmeisser et al. 2003), and could be used in future studies to determine the presence of virulence genes in drinking water microbial populations.

Including this study, nearly identical sequences have been recovered in several independent studies from our laboratory over a period of 5 years. Our results from clone libraries generated using different 16S rRNA gene primers, and using both RNA and DNA as a target, suggests the ubiquity and potential active status of these bacterial groups in the studied water distribution system. Our results also underline the importance of developing molecular-based methods that both detect and enumerate active uncultured bacteria in drinking water systems. Such methods will prove to be useful in studying the ecology and survival characteristics of these bacteria, which in turn will help us better understand potential risks associated with the common bacterial groups found in potable water. Our continuing objective is to elucidate the community dynamics in distribution water and biofilm, particularly in regard to active members of drinking water microbial communities to ascertain how this affects the survival of opportunistic pathogens present in WDS biofilm and water.

ACKNOWLEDGMENTS

Any opinions expressed in this paper are those of the authors and do not necessarily reflect the official positions and policies of the USEPA. Any mention of products or trade names does not constitute recommendation for use by the USEPA.

REFERENCES

Altshul, S.F.; Madden, T.L.; Schaffer, A.A.; Zhang, J.; Miller, W.; Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research, **25** (17), 3389.

Amann, R.I.; Lemmer, H.; Wagner, M. (1998) Monitoring the community structure of wastewater treatment plants: a comparison of old and new techniques. FEMS Microbiology Ecology, **25**, 205.

Babenzien, H-D. (1989) Genus *Neskia* Famintzin 1892 484^{AL} Bergey's Manual of Systematic Bacteriology, Vol.3 (Staley J.T.; Bryant, M.P.; Pfennig, N; Holt, J.G., eds), pp. 1979-1981. Williams & Wilkins, Baltimore, MD.

Cole J.R.; Chai B.; Marsh T.L.; Farris R.J.; Wang Q.; Kulam S.A.; Chandra S.; McGarrell D.M.; Schmidt T.M.; Garrity G.M.; Tiedje J.M. (2003) The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. Nucleic Acids Research, **31**(1), 442.

Cooksey, R.C.; de Waard, J.H.; Yakrus, M.A.; Rivera, I; Chopite, M.; Toney, S.R.; Morlock, G.P.; Butler, W.R. (2004) *Mycobacterium cosmeticum* sp. nov., a novel rapidly growing species isolated from a cosmetic infection and from a nail salon. International Journal of Systematic and Evolutionary Microbiology, **54**, 2385.

Corpe, W.A.; Jensen, T.E. (1996) The diversity of bacteria, eukaryotic cells and viruses in an oligotrophic lake. Applied Microbiology and Biotechnology, **46**, 622.

Falkingham, J.O.; Norton, C.D.; LeChevallier, M.W. (2001) Factors influencing numbers of *Mycobacterium avium*, *Mycobacterium intracellulare*, and other mycobacteria in drinking water distribution systems. Applied and Environmental Microbiology, **67**, 1225.

Holland, S. (1998) aRareFactWin. University of Georgia, Athens, GA.

Hughes, J.B.; Hellman, J.J.; Ricketts, T.H.; Bohannan, B.J.M. (2001) Counting the uncountable: statistical approaches to estimating
Microbial diversity. Applied and Environmental Microbiology, **67** (10), 4399.

Humrighouse, B.W.; Santo Domingo, J.W.; Revetta, R.P.; Lamendella, R; Kelty, C.A.; Oerther, D.B. (2006) Microbial characterization of drinking water systems receiving groundwater and

surface water as the primary sources water. In: Proceedings of 8th Annual Water Distribution Systems Analysis Symposium, Cincinnati, Ohio, August 27-30, 2006.

Hurlbert, S.H. (1971) The nonconcept of species diversity: a critique and alternative parameters. Ecology, **52**, 577.

Kalmbach, S.; Manz, W.; Wecke, J.; Szewzyk, U. (1999) *Aquabacterium* gen. nov., with description of *Aquabacterium citratiphilum* sp. nov., *Aquabacterium parvum* sp. nov. and *Aquabacterium commune* sp. nov., three in situ dominant bacterial species from the Berlin drinking water system. International Journal of Systematic Bacteriology, **49**, 769.

Keinanen, M.M.; Martikainen, L.K.; Korhonen, M.H. (2003) Microbial community structure in biofilms and water of a drinking water distribution system determined by lipid biomarkers. Water Science and Technology, **47** (5), 143.

Keinanen-Toivola , M.M.; Revetta, R.P.; Santo Domingo J.W. (2006) Identification of active bacterial communities in a model drinking water biofilm system using 16S rRNA-based clone libraries. FEMS Microbiology Letters, **257**, 182.

Lane, D. J.; Pace, B.; Olsen, G. J.; Stahl, D. A.; Sogin, M. L.; Pace, N. R. (1985) Rapid determination of 16S ribosomal sequences for phylogenetic analyses. Proceedings of the National Academy of Sciences.USA, **82**, 6955.

LeChevallier, M.W.; Welch, N.J.; Smith, D.B. (1996) Full-scale studies of factors related to coliform regrowth in drinking water. Applied and Environmental Microbiology, **62**, 2201.

Ludwig, W.; Strunk, O.; Westram, R.; Richter, L.; Meier, H.; Yadhukumar, A.B.; et. al. (2004) ARB: a software environment for sequence data. Nucleic Acids Research, **32** (4), 1363.

Martiny, A.C.; Jorgensen, T.M.; Albrechtsen, H.J.; Arvin, E.; Molin, S. (2003) Long-term succession of structure and diversity of a biofilm formed in a model drinking water distribution system. Applied and Environmental Microbiology, **69**, 6899.

Meckes, M.C.; Haught, R.C.; Dosani, M.; Clark, R.M.; Sivaganesan, M. (1999) Effect of pH adjustments on biofilm in a simulated water distribution system. In: Proceedings of the American Water Works Association Water Quality Technology Conference, October 31-November 3, Denver, AWWA.

Park, M.; Yun, S.T.; Kim, M.S.; Chun, J.; Ahn, T.I. (2004) Effect of chlorine on the bacterial colonization of a model distribution system. Revue des Sciences de L'eau, **5**, 399.

Pladdies, T.; Babenzien, H-D.; Cypionka, H. (2004) Distribution of *Nevskia ramosa* and other rosette-forming neustonic bacteria. Microbial Ecology, **47**, 218.

Pozos, N.; Scow, K.; Wuertz, S.; Darby, J. (2004) UV disinfection in a model distribution system: biofilm growth and microbial community. Water Research, **38**, 3083.

Reasoner D.J.; Geldreich, E.E. 1985. A new medium for the enumeration and subculture of bacteria from potable water. Applied and Environmental Microbiology, **49,** 1.

Saitou, N.; Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution, **4** (4), 406.

Santo Domingo, J.W.; Meckes, M.C.; Simpson, J.M.; Sloss, B.; Reasoner, D.J. (2003) Molecular characterization of bacteria inhabiting a water distribution system simulator. Water Science and Technology, **47**(5), 149.

Schmeisser, C.; Stockigt, C.; Raasch, C. (2003) Metagenome survey of biofilms in drinking-water networks. Applied and Environmental Microbiology, **69**, 7298.

Shigematsu, T.; Yumihara, K.; Ueda, Y.; Numaguchi, M.; Morimura, S.; Kida, K. (2003) *Deltia tsuruhatensis* sp. nov., a terephthalate-assimilating bacterium isolated from activated sludge. International Journal of Systematic and Evolutionary Microbiology, **53**, 1479.

Sturmeyer, H.; Overmann, J.; Babenzien, H-D.; Cypionka, H. (1998) Ecophysiological and phylogenetic studies of *Nevskia ramosa* in pure culture. Applied and Environmnetal Microbiology, **64**, 1890.

Williams, M.M.; Braun-Howland, E.B.; (2003) Growth of *Escherichia coli* in model distribution system biofilms exposed to hypochlorous acid or monochloramines. Applied and Environmental Microbiology, **69**, 5463.

Williams, M.M.; Santo Domingo, J.W.; Meckes, M.C.; Kelty, C.A.; Rochon, H.S. (2004) Phylogenetic diversity of drinking water bacteria in a distribution system simulator. Journal of Applied Microbiology, **96**, 954.

Williams, M.M.; Santo Domingo, J.W.; Meckes, M.C. (2005) Population diversity in model potable water biofilms receiving chlorine or chloramine residual. Biofouling, **21**(5/6), 1.