RESEARCH ARTICLE

Bacterial community composition in low-flowing river water with different sources of pollutants

Abasiofiok Mark Ibekwe[1], Menu B. Leddy[2], Richard M. Bold[2] & Alexandria K. Graves[3]

[1]USDA-ARS, U.S. Salinity Laboratory, Riverside, CA, USA; [2]Orange County Water District, Fountain Valley, CA, USA; and [3]Department of Soil Sciences, North Carolina State University, Raleigh, NC, USA

Correspondence: Abasiofiok Mark Ibekwe, USDA-ARS-U. S. Salinity Laboratory, 450 W. Big Springs Rd, Riverside, CA 92507, USA. Tel.: +1 951 369 4828; fax: +1 951 342 4964; e-mail: Mark.Ibekwe@ars.usda.gov

Received 17 May 2011; revised 2 September 2011; accepted 9 September 2011. Final version published online 13 October 2011.

DOI: 10.1111/j.1574-6941.2011.01205.x

Editor: Julian Marchesi

Keywords
microbial diversity; watershed; nonpoint source pollutants; terminal restriction fragment length polymorphism; temporal variations; spatial variation.

Abstract

Pollution of water resources is a major risk to human health and water quality throughout the world. The purpose of this study was to determine the influence of pollutant sources from agricultural activities, urban runoffs, and runoffs from wastewater treatment plants (WWTPs) on bacterial communities in a low-flowing river. Bacterial community structure was monitored using terminal restriction fragment length polymorphism (T-RFLP) and 16S rRNA gene clone library. The results were analyzed using nonmetric multidimensional scaling (NMDS) and UniFrac, coupled with principal coordinate analysis (PCoA) to compare diversity, abundance, community structure, and specific functional groups of bacteria in surface water affected by nonpoint sources. From all the sampling points, *Bacteria* were numerically dominated by three phyla – the *Proteobacteria*, *Bacteroidetes*, and *Cyanobacteria* – accounting for the majority of taxa detected. Overall results, using the β diversity measures UniFrac, coupled with PCoA, showed that bacterial contamination of the low-flowing river was not significantly different between agricultural activities and urban runoff.

Introduction

Numerous studies have shown that individual bacterial populations are highly dynamic and can differ strongly in their response to resource availability such as organic carbon, nitrogen, and phosphorus and to food web structure (Fisher *et al*., 1998; Langenheder & Jurgens, 2001; Schafer *et al*., 2001; Kent *et al*., 2004). Others have shown that shifts in microbial community structure can be related to seasonal cycles in the source water and dissolved organic matter (Sommer *et al*., 1986; Crump *et al*., 2003). Similarly, seasonal shifts in water column stability and water temperature may demonstrate an annual pattern of bacterial community variability (Murray *et al*., 1998). Some physical and chemical factors may result in changes in bacterial composition in the water column affecting the overall quality of the water. Determining physical and chemical factors, such as salinity, temperature, pH, geography, etc., that correlate with differences between diverse microbial communities will reveal how microorganisms tolerate different kinds of environmental change and increase our understanding of microbial ecology and evolution and their effects on public health. In addition, determining the environment types that contain the most phylogenetic diversity will reveal where new sequencing efforts to catalog global bacterial diversity will be most efficient at uncovering deep branching lineages (Lozupone & Knight, 2007).

There have been no studies in the Santa Ana River (SAR) on the influence of different inputs from the region on microbial community structure, although studies have been conducted to determine the effects of different pollutant sources on fecal indicator bacteria (Izbicki *et al*., 2002). This is an effluent-receiving river with seasonal flows owing to discharges from wastewater treatment plants (WWTPs) and flows from urban and agricultural runoffs. Owing to these activities, the quality of the river is degrading as a consequence of excess load of pollutants including microorganisms. This is due in part to nonpoint and agricultural runoff and discharges

© 2011 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved

FEMS MICROBIOLOGY ECOLOGY

from the WWTPs and urban runoff. Also, the SAR is exposed to seasonal changes in its biological, chemical, and physical environments as a result of both biotic and abiotic factors. It is, therefore, hypothesized that these changes have a significant influence on the microbial community structure of the river. Originally, the region was developed as an agricultural area, with approximately 130 km^2 of Dairy Preserve, which was one of the largest concentrations of dairies in the United States, but in recent years, the area has experienced major development and rapid increases in population. Land use in the area ranges from dense urban development to undeveloped wilderness. The quality of surface water varies throughout the SAR watershed. The surface water flowing out of the surrounding mountain is of good water quality, as water progress to the SAR, the quality deteriorates (Izbicki *et al.*, 2002; Rice, 2005). The SAR is critical for replenishment of Orange County's groundwater basin as over 2 million residents in Orange County depend on groundwater for 75% of their water supply. Any factor in the watershed which degrades the river affects the drinking water supply. The river extends from its headwaters in the San Bernardino Mountains into the Prado Basin and Santa Ana Canyon. Below Prado Dam, there are extensive facilities to recharge much of the flows in the River into the underlying groundwater basin.

To address the changes in bacterial community within the watershed, we performed a comparative study on the bacterioplankton in the surface waters by terminal restriction fragment length polymorphism (T-RFLP) of polymerase chain reaction (PCR)-amplified 16S rRNA genes (Liu *et al.*, 1997; Dunbar *et al.*, 2000, 2001; Blackwood *et al.*, 2003; LaMontagne *et al.*, 2003) and cloning and sequencing of 16S rRNA genes. Total bacteria counts using epifluorescence microscopy were used to obtain quantitative data from surface water that can relate to some environmental variables. The aims of this study were to determine the influence of pollutant source, agricultural activities from concentrated animal feeding operations (CAFOs), runoffs from residential and industrial activities, and runoffs from WWTPs, on bacterial communities in a low-flowing river during different seasons and to determine the influence of some water chemistry parameters on microbial community composition and structure.

Materials and methods

Site description and water sampling

The SAR watershed is located approximately 60 km southeast of Los Angeles, CA (Fig. 1). Urban runoff from Chino creek and agricultural runoff from Cypress channel discharge into the middle Santa Ana River (MSAR). The mean annual rainfall for SAR watershed is $\leq$ 800 mm per annum, predominantly between December and April. As such, the base stream flow is highly variable between seasons. The mean annual stream flow from USGS-gauged data from Chino Creek (S 3-Chino creek @ Schaefer Ave) was 133.6 $m^3\ s^{-1}$ and at Cypress channel (S6 – Cypress channel @ Schaefer Ave) was 96.8 $m^3\ s^{-1}$. Sampling sites used for this study are shown in Table 1; Fig. 1. Locations were selected for surface water sample

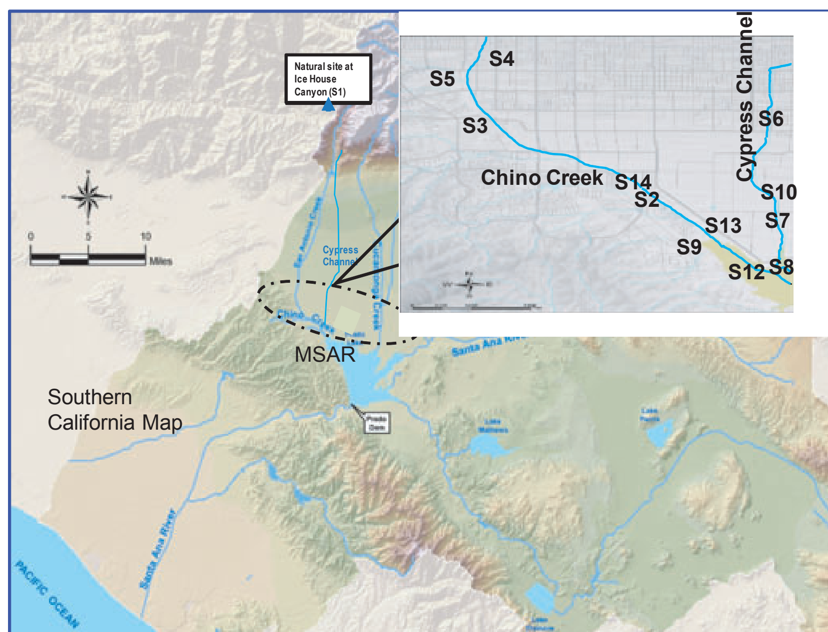


Fig. 1. Various sampling points along Chino creek and Cypress channel within the MSAR watershed. Water flow from the natural site at Ice house Canyon (S1) to the San Antonio creek and into Chino creek. This flows into the Prado basin and into Santa Ana River and finally empties into the Pacific Ocean. The Santa Ana River is critical for the replenishment of Orange County's Groundwater Basin as over 2 million residents in Orange County depend on groundwater for 75% of their water supply.

© 2011 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved

Table 1. Sampling locations for source affect bacterial community structure

Site #	Site locations	Land use	GPS units
S1	Ice House Canyon	Open space/natural site	N34^0 15.057 min; W117^0 37.977 min; 1447-m elevation
S2	Chino Creek @ Central avenue	Urban runoff	N33^0 58.420 min; W117^0 41.302 min; 174-m elevation
S3	Chino Creek @ Schaefer avenue	Urban runoff	N34^0 00.246 min; W117^0 43.628 min; 207-m elevation
S4	San Antonio Wash @ County drive	Urban runoff + commercial wash out	N30^0 01.543 min; W117^0 43.652 min; 223-m elevation;
S5	Chino Creek @ Riverside drive	Urban runoff	N34^0 01.144 min; W117^0 44.204 min; 217-m elevation;
S6	Cypress Channel @ Schaefer avenue	Agricultural	N34^0 00.262 min; W117^0 39.766 min 209-m elevation;
S7	Cypress Channel @ Kimball avenue.	Agricultural	N33^0 58.113 min; W117^0 39.624 min 177-m elevation;
S8	Cypress Channel @ Golf course	Agricultural	N33^0 57.057 min; W117^0 39.555 min; 160-m elevation;
S9	Big League Dreams storm drain	Urban runoff + possible agricultural runoff during storm events	N33^0 57.364 min; W117^0 40.788 min; 163-m elevation;
S10	Dirt channel on Kimball	Agricultural	N33^0 58.109 min; W117^0 40.286 min 184-m elevation;
S12	Chino Creek @ Pine avenue	Urban runoff + wastewater	N33^0 56.941 min; W117^0 39.986 min; 155-m elevation;
S13	WWTP	Effluent from WWTP	N33^0 57.840 min; W117^0 40.826 min; 180-m elevation;
S14	WWTP	Effluent from WWTP	N33^0 58.799 min; W117^0 41.655 min; 184-m elevation;

analyses based on historical data obtained for the Total Maximum Daily Loads (TMDL) for Bacterial Indicators for MSAR watershed. All sampling locations with site names, descriptions, and geographic positioning system (GPS) coordinates are listed in Table 1. Water samples at two WWTPs were retrieved from the sampling ports located at the treatment plant site for sample collection (Table 1). The plants discharged tertiary-level-treated water downstream resulting in continuous but variable stream flow throughout the year along Chino creek. Cypress channel is more affected by dairy or agricultural runoff and Chino creek affected more by WWTP and urban runoff. The Ice House Canyon (S1; Table 1) is a natural site and was used mainly as the control sites because runoff from this site was mainly from melting snow with little or no contamination. Ice House Canyon Creek is located in the San Gabriel Mountains and is a tributary to San Antonio Creek approximately 2.1 km upstream of Mt. Baldy Village. Historical data for Ice House Canyon for fecal coliform had averaged 9 CFU 100 mL^{-1} over a 5-year time period covering from 2000 to 2005 (Rice, 2005).

Reference water samples were taken under storms, recessional, and dry-weather flows. A water sample was considered a ‘storm flow’ sample if there was > 12.7 mm of accumulated precipitation (i.e. rain). In situations where the storm continued for several days to a week, every attempt was made to retrieve the specific ‘storm-flow’ samples for analyses at the peak of the storm. Three separate storm weather events were sampled during the study. The largest of the three storms was 30 mm of precipitation in the area, followed by another storm with 25 mm of average precipitation and 21 mm of average precipitation. Recessional flow samples were collected approximately 72 h following a storm event. Dry-weather flows were analyzed during the summer months between May and August when no precipitation was recorded in the area between 30 and 90 days. Seventy-six storm-flow samples were collected for the study. A total of 112 recessional flow samples and 90 dry-weather-flow samples were collected for the study, and an accumulated total of 278 water samples.

For sample collection, sterile Nalgene sampling bottles were used (APHA, 1995). All samples were collected in duplicates. For sites that were deep enough to obtain sample, grab samples were collected about 10–15 cm below the surface of the water. Sites with shallow flow were sampled using sterile stainless steel sampling device. Field parameters consisting of electrical conductivity, pH, temperature, turbidity, and dissolved oxygen (DO) were taken at each sample location. All water samples were transported on ice to the laboratory and analyzed within 6 h. Sample turbidity was determined using a Hach model 2100P Portable Turbidimeter (Loveland, CO) according to manufacturer’s instructions and calibrated each day of use.

Total bacteria in reference water samples were collected by filtration onto a 0.2-micron black polycarbonate, 25-mm membrane filter (Millipore isopore membrane filters, GTBP), placed on a 125-mL Erlenmeyer filter flask that contains a 25-mm glass microanalysis vacuum filter holder and support. The water sample was filtered and stained with a water-soluble DNA-binding fluorochrome, 4′, 6-diamidino-2-phenylindole (DAPI). The sample was held on the vacuum filter holder for approximately 10 min, after which the filter was removed and placed on a glass slide with a cover slip. The glass slide was examined under the microscope and the cells, which fluoresce

© 2011 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved

blue under 350- to 380-nm illumination, counted. The bacteria were counted in 30 individual fields with the 100× oil emersion lens using epifluorescence microscopy (Porter & Feig, 1980; Hobbie *et al.*, 1997), and the average number of bacteria per milliliter was calculated.

DNA extraction, T-RFLP, and sequencing analysis

Total bacterial DNA was extracted from a 250-mg pellet of concentrated water sample (1 L) centrifuged at 9000 ***g*** for 30 min. DNA was extracted using UltraClean water DNA kits (MO BIO, Inc., Solana Beach, CA) according to the manufacturer's protocol with slight modifications and stored at −20 °C. DNA was visualized by agarose gel electrophoresis using the 4% NuSieve 3 : 1 agarose reliant gel system (cambrex biosciences, Rockland, ME) and quantified by UV spectrophotometry using the CCD array UV-VIS spectrophotometer (Spectral Instruments, Inc.). 16S rRNA genes from purified DNA samples were PCR-amplified using universal bacterial primers 6 HEX (fluorescently labeled forward primer), F 5′ CGGCAGGCCTA ACACATGCAAGTCG 3′ and reverse primer 5′ GGTTGC GGCCGTACTCCCCAGG 3′ (Avaniss-Aghajani *et al.*, 1994). All primers were synthesized and labeled by IDT (Coraville, IA). PCR amplification was carried out using approximately 40 ng of total DNA for each sample. Each 100-μL PCR mixture contained 1× PCR buffer with 1.5 mM $MgCl_2$ (Applied Biosystems, ABI, Foster City, CA), 200 μM each deoxynucleotide triphosphate, 0.2 μM of each primer, and 2.5 U of AmpliTaq gold (ABI). The forward primer was labeled with 5-hexachlorofluorescein (HEX) and run with a GeneAmp PCR system 9600 (ABI) with the following conditions: 10 min hot start at 95 °C and then 35 cycles consisting of 94 °C for 1 min, 58 °C for 1 min, and 72 °C for 3 min, followed by final 10-min extension at 72 °C. Amplicons were visualized for quality control purposes on a 4% NuSieve 3 : 1 agarose gel system (Cambrex BioSciences, Rockland, ME.) with a size ladder to approximate the size of the amplicon. For each PCR, positive and negative controls were included for quality control and verification of PCR performance.

All T-RFLP analyses were carried out as previously described (Dunbar *et al.*, 2000, 2001; Blackwood *et al.*, 2003; LaMontagne *et al.*, 2003). In brief, all fluorescently labeled amplicons were digested with DdeI and MspI enzymes (New England Biolabs, Beverly, MA) at 37 °C for 3 h (Avaniss-Aghajani *et al.*, 1994). Each 20-μL digestion reaction contained approximately 40 ng of PCR products and 10 U of DdeI and MspI enzymes in separate reactions at 37 °C for 3 h. Fluorescently labeled terminal fragments were separated by electrophoresis and fragment lengths determined with an ABI 310 Genetic Analyzer as follows: 2 μL of the amplicons or the fluorescently labeled terminal fragments were mixed with 12.5 μL of a mixture of deionized formamide (ABI) and Tamara 500 internal size standard (ABI). The mixture was denatured at 94 °C for 5 min and immediately chilled on ice prior to electrophoresis with the automated DNA sequencer in the GeneScan mode. The retention times, peak height, size (in base pairs), peak area, and data point were calculated using the Tamara internal size standard (ABI) and using the Local Southern algorithm in GeneScan (ABI) software, version 2.1 (ABI). Labeled terminal restricted fragments were aligned manually to identify peaks shared between profiles generated from replicates and reference samples; fragments that differed by < 1 base pair in size were considered shared. All peaks were analyzed, and peak size and total area were recorded in Excel for further analysis. After acquiring the restriction fragment patterns using Genescan software, fragment lengths and peak areas were compared with fragment lengths and peak areas from other locations using Microsoft Excel, and the data were used for multivariate analysis.

The same PCR conditions used for terminal restriction fragment length polymorphism (T-RFLP) analysis were used for 16S rRNA gene cloning and sequencing, with the exception that the forward primer was not labeled. Cloning was performed with the TOPO TA Cloning Kit (Invitrogen, Rockville, MD) for sequencing according to the manufacturer's instructions. Partial sequences with an average length above 600 bp were obtained using the forward primer for amplification reactions in an ABI 3700 system analyzer.

Data analysis

T-RFLP profiles were analyzed using the IBEST tools (Abdo *et al.*, 2006) and T-Align software (Smith *et al.*, 2005). Data from DdeI and MspI were combined and ordered by nonmetric multidimensional scaling (NMDS) with Wisconsin transformation and Bray–Curtis index, and correlations between the ordination axes and environmental variables were calculated in the PC-ORD version 5 (McCune & Mefford, 1999) for sample ordination purposes. Regression analysis was conducted in SAS (SAS, 2005) to determine the associations between total bacterial population measured by an epifluorescence microscopy and environmental variable via the method of stepwise model selection. Five additional field parameters, pH, salinity (EC), DO, turbidity, and temperature of the surface water were also measured at each sample point. Variations in each of these water quality parameters are known to affect bacterial concentrations. Therefore, the field parameters were measured as possible covariate in the subsequent statistical analyses using analysis of

© 2011 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved

covariance (ANOCOVA) (Montgomery, 2001). Tukey HSD test was used to test for significant differences in environmental variables among the different sites. The comparison of the diversity was made using a one-way analysis of variance, and Tukey HSD test for post hoc analysis (SAS, 2005).

Sequences from the clone libraries were screened for chimeras using Bellerophon (Huber *et al.*, 2004), and potential chimeras were excluded from further analysis. Distance matrices were constructed for clone sequence libraries using the dist.seqs function in MOTHUR, version 1.9.1 (Schloss *et al.*, 2009). MOTHUR was also used to assign sequences to operational taxonomic units (OTUs, 97% similarity) and calculate both Shannon's diversity index values (H), ACE and ChaoI richness estimates. Principal coordinate analysis (PCoA) and hierarchical clustering in UniFrac were carried out using MOTHUR. PCoA is similar to principal component analysis, except that the starting point is a matrix of distances between samples rather than a matrix of observations about each sample. Phylogenetic trees were constructed using the relaxed neighbor-joining algorithm in CLEARCUT (version 1.0.9) (Sheneman *et al.*, 2006), and between-site comparisons of phylogenetic structure were conducted using the parsimony test in Treeclimber (Schloss & Handelsman, 2006). Parsimony test scores with P-values < 0.05 were considered to represent significant differences, and pairwise comparisons of individual libraries were conducted only if the study-wide (that is global) test was found to be statistically significant (Schloss & Handelsman, 2006). Bonferroni correction for multiple comparisons was used among the pairwise comparisons, adjusting our significance level to $P \leq 0.0017$ (Neter, 1996; Hollister *et al.*, 2010).

Results

Environmental variables and total bacteria using epifluorescence microscopy

Physical and chemical characteristics such as temperature, turbidity, salinity (EC), DO, and pH were determined in various surface water samples retrieved during storm, recessional, and dry weather. Salinity, pH, temperature, turbidity, DO, and total bacterial data are summarized and mean data for storm, recessional, and dry weather presented (Table 2). The temperature values for surface water samples and wastewater effluent stream during storm events ranged from 6 °C at Ice House Canyon (S1) to 25 °C at WWTP (S13, S14). Electrical conductivity (EC) or salinity, pH, and DO did not show a significant fluctuation as was observed with turbidity and temperature during the study (Table 3). During three (3) separate storm sampling events, turbidity values for surface water samples varied within each storm event and were generally higher from urban and agricultural runoff than at the WWTPs and at the Ice House Canyon (S1). Average turbidity values were below 1 Nephelometric Turbidity Units (NTU) for the Ice House Canyon (S1) site. The most stable values (1.04 NTU) were observed with wastewater effluent samples (S13 and S14).

ANOCOVA results for the water quality variables for the two creeks, natural site, and WWTP suggested potentially important differences among the variables that may influence total bacterial counts dynamics. The primary parameter F and t-tests for the estimated mixed linear ANOCOVA model are shown in Table 3. As shown, the site effect was significantly below the 0.05 level in the total bacteria model. Also, the weather conditions significantly affected total bacteria ($P = 0.001$). The measured log turbidity levels were significantly related to the log bacteria count

Table 2. Mean concentrations of environmental parameters and total bacterial concentrations as determined by epifluorescence microscopy measured from all sampling sites

Variables	Dry weather	Recessional weather	Storm weather
Turbidity (NTU)	5.19 ± 0.71	6.23 ± 1.14	218.75 ± 63.77
Salinity (EC) (dS m^{-1})	0.85 ± 0.37	1.03 ± 0.68	0.49 ± 0.34
pH	8.24 ± 0.97	8.22 ± 0.99	7.82 ± 0.55
Temperature (°C)	25.70 ± 5.31	18.07 ± 5.62	15.46 ± 4.31
DO	8.13 ± 2.57	10.61 ± 3.71	9.53 ± 3.84
Total bacteria (cell mL^{-1})	8.09 ± 1.06	6.92 ± 1.45	8.47 ± 0.73

NTU, Salinity (EC) (dS m^{-1}).

Table 3. Climate and water quality trends estimated with mixed linear ANOCOVA models. All P-values computed using the Kenward–Roger adjustment technique

Factor or covariate	Statistics	Total bacteria
Site	F score	2.13
	P-value	0.034
Weather conditions	F score	8.28
	P-value	0.001
ln[Turbidity]	t score	7.90
	P-value	< 0.001
ln[ECe]	t score	−1.55
	P-value	0.128
pH	t score	−1.10
	P-value	0.275
Temperature	t score	3.46
	P-value	0.001
DO	t score	−1.251
	P-value	0.13

© 2011 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved

data, and higher turbidity level corresponded with significantly higher total bacterial counts ($P \leq 0.001$). In contrast, the log salinity levels (In[ECe] were never significantly related to total bacteria count level. In this ANOCOVA model, higher pH levels corresponded with lower bacterial counts. Also, higher water temperatures resulted in higher total bacterial counts ($P = 0.001$).

During storm events, the concentrations of total bacteria were significantly higher ($P < 0.05$) in some sites with urban runoff (S9 and S12) than from agricultural runoff (S6, S7, S8; Table 4). Also, lower numbers were observed in the effluent from WWTPs (S13, S14) in comparison with the natural, urban, and agricultural runoffs. However, during the recessional flow, total bacteria counts were not significantly different between agricultural and urban runoffs. Also, no significant differences were found during the dry-weather sampling event between urban runoff and agricultural runoff.

Comparison of bacterial community T-RFLP profiles from surface water samples

The structure of bacterial communities derived from analysis of the T-RFLP profiles from various water samples was determined using NMDS. The NMDS in Fig. 2a shows the relationships between the profiles obtained from dry-weather water samples indicating that some bacterial communities differed in structure and that these differences were related to time of sampling but not source and water quality attributes, especially those linked to salinity, temperature, and DO. The influence of salinity on microbial composition was determined to be the most significant by Monte Carlo permutation test for dry weather based on time but not on the source. The effect of time between sampling events was evidenced by clustering of bacterial communities based on sampling date. Most of the samples collected on July 2005 were grouped on the top left of the NMDS, and samples collected in June 2006 were on the bottom right of the NMDS. Therefore, four distinct clusters were formed temporally. While sampling sites were clustered temporally during the dry weather, clustering during storm conditions varied (Fig. 2b). During the storm, there was a strong influence of temperature and DO and a moderate influence of turbidity on microbial community structure. Temporal shifts affected the communities the most, while spatial clustering produced a moderate effect on some sites. Most samples from October 2005 were separated to the right of axis 1 while samples from February were to the left of axis 2. The last cluster had samples from the three separate months, which separated to the lower side of axis 1. Samples from Ice House Canyon site (S1) did not group with other samples as well as samples from agricultural runoff (S6, S7, S8, and S10) collected during December 2004. For recessional flow, four factors had the major influences on microbial community (Fig. 2c): turbidity, temperature, DO, and salinity. However, a permutation test also showed no significance influence of these variables at the 5% level ($P < 0.792$) on microbial communities. During the recessional flow, samples were spatially separated. All samples from Ice House Canyon (S1) clustered together.

To determine whether microbial community structure was strongly related to environmental variables over random relationships, NMDS ordination of sites in environmental space was evaluated. The cumulative percentage variance of species–environment correlations indicated an overall variance in the dataset for dry, storm, and recessional weather (Table 5). For dry weather, the

Table 4. Total bacteria concentrations (cell mL^{-1}) on surface water using epifluorescence microcopy method

Sites*	Storm 1	Storm 2	Storm 3	Rec 1	Rec 2	Rec 3	Rec 4	Rec5	Dry 1	Dry 2
S1	7.81bcd	7.10f	7.88g	2.70b	6.81a	6.47a	6.93e	–	6.41a	7.13f
S2	–	8.38c	8.56ef	7.70b	7.73a	7.81a	7.58abc	7.26bc	5.31a	8.34c
S3	8.65b	8.47bc	9.0cde	6.88ab	6.42a	7.78a	7.13de	6.99d	8.65a	8.56bc
S4	8.77b	8.37c	9.11cd	6.91ab	2.69b	8.15a	7.90ab	7.94ab	7.28a	8.49c
S5	8.41bc	8.38c	8.50f	5.14ab	6.27a	7.05a	7.12de	7.35bcd	8.08a	8.37c
S6	8.42bc	7.95d	–	7.07ab	–	–	–	–	8.06a	–
S7	8.19bcd	8.25cd	8.85def	7.32ab	7.25a	8.34a	8.23a	8.22a	9.01a	9.25a
S8	8.50bc	8.25cd	9.65b	8.65a	7.49a	8.90a	8.02ab	7.67abc	9.03a	9.18a
S9	9.90a	8.74b	9.36bc	4.75ab	7.22a	8.16a	–	–	8.82a	8.47b
S10	8.86b	8.47bc	7.88g	–	7.92a	–	–	–	–	–
S12	8.50bc	9.22a	10.31a	6.81ab	7.86a	754a	7.49bcd	7.33bcd	7.67a	7.96de
S13	7.22d	7.46e	7.56g	4.61ab	4.67ab	6.79a	7.26cde	7.27bcd	7.53a	7.83e
S14	7.51d	7.59e	7.94g	6.48b	7.07a	5.25a	7.68bc	6.97d	7.64a	8.26cd

–, samples not collected from the sites on those days.

*Urban runoff along Chino Creek (S2, S3, S4, S5, S9, and S12), wastewater treatment plant – WWTPs (S13, S14), agricultural runoff along Cypress channel (S6, S7, S8, S10), and natural site (S1). Means with the same letters are not significantly different from each other within each column.

© 2011 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved

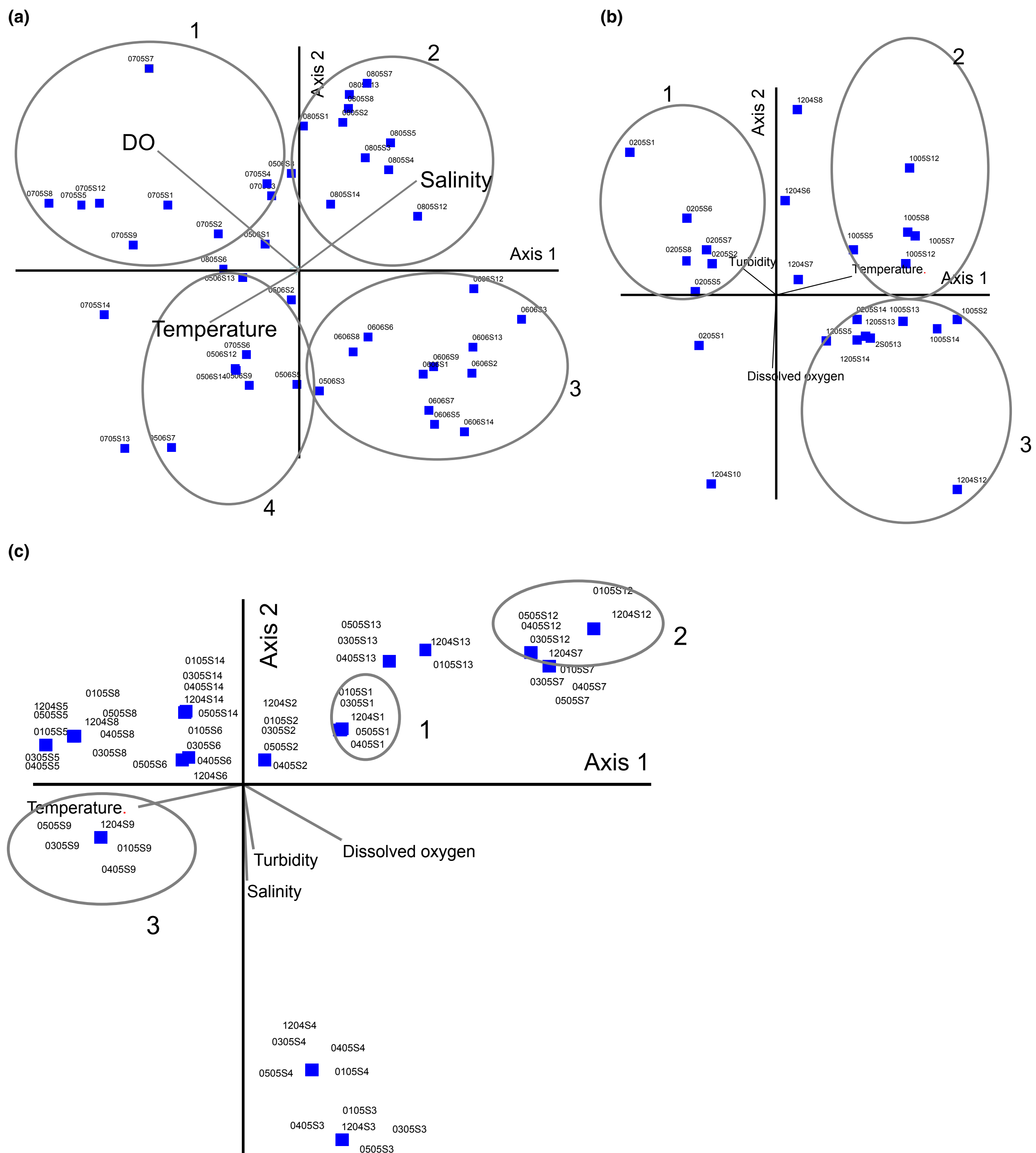


Fig. 2. Influence of environmental variables on bacterioplankton communities in the low-flowing river during (a) dry weather condition. T-RFLP identification profiles are presented with the first two numbers indicating month (07), and the second group of numbers indicating year (05) that samples were collected, and the last set of numbers beginning with letter 'S' indicating sites (S7) that samples were collected from 0705S7. In short, this sample was collected on July 2005 from site S7. The four ovals represent four clusters that samples were separated based on time. Samples from cluster 1 were taken in July 2005, cluster 2 in August 2005, cluster 3 in June 2006, and cluster 4 in May 2006. (b) Storm weather conditions with three temporal clusters 1, 2, and 3. Clusters 1 and 2 represent samples from February 2005 and October 2005, while cluster 3 has mixed samples; therefore, temporal shifts affected the communities the most, while spatial clustering produced a moderate effect on some sites. (c) During the recessional flow, samples were spatially separated. All samples from Ice House Canyon (S1) clustered together.

© 2011 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved

first axis explained 49.3% of the total variation, the second axis explained 54.1%, and the third axis explained 67.2%. Species–environment correlations for dry weather were high for all three axes (0.927, 0.834, and 0.790), indicating a significant high relationship between species (microbial community) and environmental variables. The same high levels of species–environment correlations for the three axes during storm sampling were determined (0.942, 0.962, and 0.948), also indicating a relationship between microbial communities and environmental variables. The first axis explained 73.3% of the total variation, the second axis explained 76.0%, and third axis explained 82.0%. However, this observation was low during recessional flow; the first axis explained 12.7% of the total variation, the second axis explained 20.2%, and third axis explained 27.2%, suggesting a low relationship between microbial community and environmental variables.

Taxonomic composition of surface water microbial communities

We recovered 310 partial sequences of the 16S rRNA genes from urban, agricultural, natural site, and WWTPs. Sequences were assigned to the following bacterial groups: *Acidobacteria* (0.65%), *Bacteroidetes* (17.10%), *Cyanobacteria* (20.65%), *Firmicutes* (0.97%), *Proteobacteria* (53.23%), including representatives of the classes *Alpha, Beta, Delta* and *Gamma Proteobacteria* and other smaller phyla (Fig. 3). Close examination of the different sources showed that Cypress channel (agricultural source) has the highest percent sequences for *Bacteroidetes* and *Cyanobacteria,* while urban runoff from Chino creek, WWTPs, and natural site was dominated by *Proteobacteria* sequences. Rarefaction curves (Fig. 4) and diversity indices (Table 6) were determined for the four sources. The linear rarefaction curves provided evidence that the bacterial diversity is far from saturation at the high similarity level (97%, $\mathrm{OTUs}_{0.03}$). None of the curves reaches a plateau. This result is confirmed by the comparison of the observed (rarefaction index) and estimated (nonparametric ACE-abundance base coverage estimator and Chao1 indices) OTUs at 97% level of similarity (Table 6). The ACE and the Chao1 richness values were relatively far from the observed ones (rarefaction), confirming that our sampling was not close to saturation. Examination of the rarefaction values between 50% and 80% showed no stability in values, suggesting that much more sequences must be analyzed to reach a plateau (data not shown).

Phylogenetic structure of bacterial community from different zones

Samples from the different sources on the distribution of bacterial phylogenetic similarity were sorted into different groups by applying PCoA (Fig. 5), and the UPGMA hierarchical clustering analysis (data not shown) to a matrix of UniFrac distances using the UniFrac web interface in MOTHUR. Source classification was a strong structural factor of the bacterial assemblages (R^2 = 0.32, P = 0.001), and bacteria grouped according to the source of origin. Inputs from urban runoff from Chino creek and agricultural runoff from Cypress channel were not significantly different from each other (P = 0.08), but the two sources were significantly different (P = 0.001) from the natural site and from WWTPs (Fig. 5). The PCoA (Fig. 5) showed samples from urban runoff and agricultural runoff (Cypress channel) clustering to the bottom left, while samples from WWTPs clustered to the top left and the natural site to the bottom right. This was confirmed by hierarchical clustering analysis (data not shown) with Jackknife supporting values, which showed that urban and agricultural runoffs were not different from each other but were different from the natural site and WWTPs bacterial community structure.

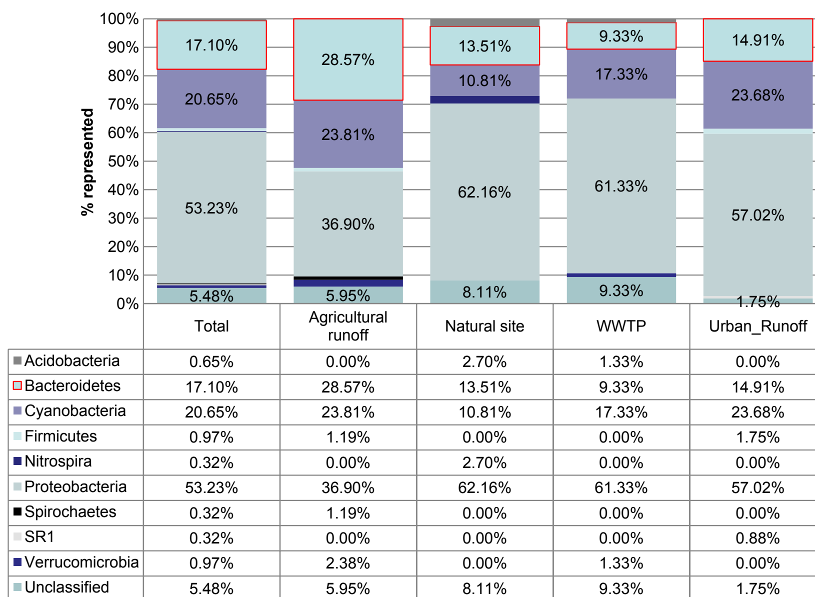
A comparison of the community structures detected from the four sources is reported using the Venn diagram with the number of shared OTUs among the different sources (data not shown). The majority of $\mathrm{OTUs}_{0.03}$ rep-

Table 5. Eigenvalues and variance decomposition for weather conditions

Weather condition	Axes	Eigenvalues	Species–environment correlations	P*	Cumulative% variance of species–environment correlations
Dry	1	0.640	0.927	0.001	49.3
	2	0.321	0.834		54.1
	3	0.267	0.790		67.2
Storm	1	0.443	0.942	0.001	73.3
	2	0.364	0.962		76.0
	3	0.353	0.948		82.0
Recessional	1	0.136	0.549	0.792	12.7
	2	0.082	0.467		20.2
	3	0.047	0.312		27.2

*P is not reported for axes 2 and 3 because using a simple randomization test for these axes may bias the P-value.

© 2011 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved



	Total	Agricultural runoff	Natural site	WWTP	Urban_Runoff
Acidobacteria	0.65%	0.00%	2.70%	1.33%	0.00%
Bacteroidetes	17.10%	28.57%	13.51%	9.33%	14.91%
Cyanobacteria	20.65%	23.81%	10.81%	17.33%	23.68%
Firmicutes	0.97%	1.19%	0.00%	0.00%	1.75%
Nitrospira	0.32%	0.00%	2.70%	0.00%	0.00%
Proteobacteria	53.23%	36.90%	62.16%	61.33%	57.02%
Spirochaetes	0.32%	1.19%	0.00%	0.00%	0.00%
SR1	0.32%	0.00%	0.00%	0.00%	0.88%
Verrucomicrobia	0.97%	2.38%	0.00%	1.33%	0.00%
Unclassified	5.48%	5.95%	8.11%	9.33%	1.75%

Fig. 3. Surface water bacterial taxonomic composition by *Phyla* from the different sources. Clones used for the table are combined from the three seasons. Table below the graph is an expanded legend to show percentages as some phyla had very small percentages. Taxa represented occur at $\geq$ 0.1% abundance in at least one sample.

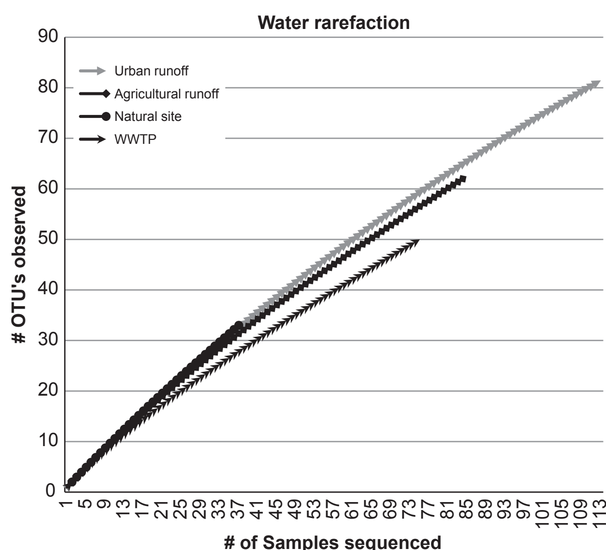


Fig. 4. Rarefaction curves of OTUs defined by 3% sequence variation in urban runoff, agricultural runoff, natural site, and WWTPs.

resented community members that are specific to urban runoff (82) agricultural runoff (62), natural site (33), and WWTPS (50) (Table 6) microbial communities. As expected, the natural site and urban runoff microbial communities had few OTUs in common (i.e. two shared OTUs). In contrast, the overlap for urban runoff and agricultural runoff (Cypress channel) shared 16 OTUs, while WWTPs and control shared 3. We did not do further comparisons because of the small numbers of OTUs in this dataset.

Table 6. The ability of the two nonparametric richness estimators (ACE and Chao1) to predict the number of bacterial OTUs from different zones compared with the numbers of observed OTUs (rarefaction index) at 97% level

Zones	Rarefaction	ACE	Chao1
Urban runoff	82	202	196
Agricultural runoff	62	241	237
Natural site	33	197	178
WWTPs	50	212	265

Data are computed from clones isolated from samples collected during storm, dry weather, and recessional flow combined.

Discussion

The association between environmental variables, sources of pollutants, and bacterial communities in low-flowing rivers in an urban environment is still not well understood. This river is highly impacted by inputs from agricultural activities and urban runoff. In this study, we found salinity to have the strongest influence on bacterial community structure during dry weather. Bernhard *et al.* (2005) showed similar findings as in our study and noted that salinity was one of the primary factors influencing microbial communities in freshwater environments with a low flow. Also, similar results were found along a salinity gradient in Plum Island Sound, MA, using denaturing gradient gel electrophoresis (DGGE) analysis (Crump *et al.*, 2004). The findings from our study under low flow conditions and in other geographic areas such as estuaries suggest a predictable effect of salinity on low flow rivers on microbial communities which may be universal. In a

© 2011 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved

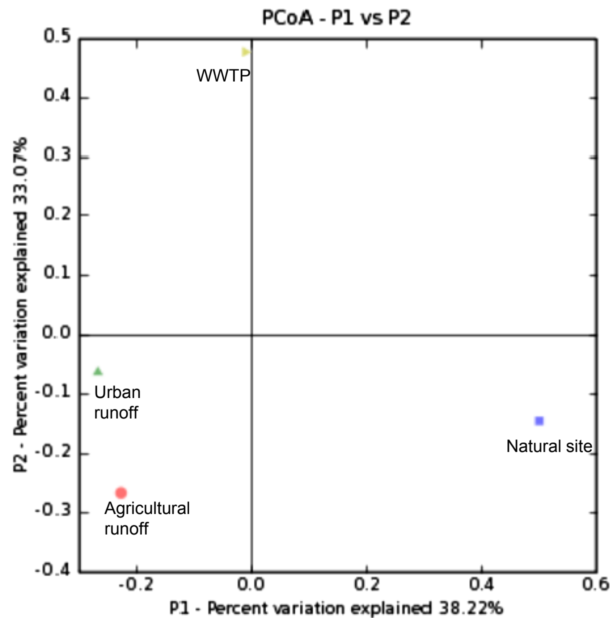


Fig. 5. PCoA obtained with the UniFrac distance matrix comparing the four zones. Principal coordinate 1 (P1) vs. principal coordinate 2 (P2) are represented.

study of global patterns in bacterial diversity that reported one of the most comprehensive analysis of the environmental distribution of bacteria, based on 21 752 16S rRNA gene sequences compiled from 111 studies of diverse physical environments, it was found that the major environmental determinant of microbial community composition was salinity rather than extremes of temperature, pH, or other physical and chemical factors represented in the samples based on similarities in the phylogenetic lineages (Lozupone & Knight, 2007). Temperature was another factor that was closely correlated with bacterial communities during the different seasons (Fig. 2a–c). This is in agreement with many other studies that correlate the influence of environmental factors with bacterial community structure in a freshwater basin (Gerdts *et al.*, 2004; Kent *et al.*, 2004; Bernhard *et al.*, 2005; Crump & Hobbie, 2005; Sapp *et al.*, 2007;). Another significant factor that correlated with bacterial communities during the various seasons was DO.

The impact of other variables such as turbidity was only important during the storm and recessional flow. This suggests that during storm, there was an increase in nutrient load carried along the channels from different sources, which may result in bacterial growth in water samples or transport of bacteria from areas of high concentration to areas of low concentration. While salinity was a significant factor during dry weather, turbidity was a moderate factor influencing bacterial communities during the storm and recessional flow, according to the intraset correlation (Fig. 2b and c). Storm effect is clearly explained by the February 2005 storm, which was the largest of the three storms with 30 mm of precipitation in the area. In some recent studies, it was shown that a strong correlation between the total fecal load (kg day^{-1}), bovine fecal loads (kg day^{-1}), *Escherichia coli* load rate (CFU day^{-1}), 7-day antecedent precipitation, and turbidity in a stream with a mixed land-use watershed (Gentry *et al.*, 2006, 2007) was very common. These authors used various datasets to establish parameter correlations and spatial dependencies throughout the watershed. They concluded that source-specific bacterial loads can be used to understand hydrologic influences on bacterial delivery and persistence in the system. In our study, source-specific loads correlated with microbial community structure during the recessional flow and partially during storm events. This was not the case during dry-weather conditions. It should be noted that microorganisms in rivers are diverse and dynamic in composition owing to environmental stresses (Nogales *et al.*, 2007) and therefore the composition of a microbial community in a river has been suggested as an indicator for pollution (Atlas, 1984).

Another unexpected result was the lack of clear evidence of bacterial community structure differences between the sites with agricultural activities and those receiving urban runoff (Fig. 5). Our study is in agreement with the study of Crump & Hobbie (2005), which showed the bacterial community composition of two nonintersecting temperate rivers was nearly identical and changed synchronously over 2.5 years, suggesting that intrinsic controls on bacteria were similar in the two rivers and that seasonal changes were driven by extrinsic factors such as climate. The authors showed also that temperature and river flow rate were the best predictors of temporal patterns in community structure. Another study by Stepanauskas *et al.* (2003) characterized bacterial communities in the Sacramento–San Joaquin river delta, including samples from rivers upstream of the delta. They found that these rivers contained similar communities that shifted consistently with the seasons. Similar results have been found on population structure, persistence, and seasonality of autochthonous *E. coli* in temperate, coastal forest soil from a Great Lakes watershed (Byappanahalli *et al.*, 2003, 2006). In contrast, Masin *et al.* (Masin *et al.*, 2003) showed that two rivers in the same part of the Czech Republic contained statistically different communities based on cell counts using *in situ* hybridization of the phylum- and subphylum-specific probes.

Conclusions

Using molecular methods, we were able to understand some of the relationships between microbial community

© 2011 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved

and environmental variables during different seasons and storm conditions at different sources in a low-flowing river. By studying a low-flowing river in an urban area that is impacted by more than 1.4 million people and over 200 000 cattle, we were able to relate diversity to environmental factors such as salinity, temperature, DO, and turbidity and demonstrated that bacteria can respond differently to chemical and physical parameters. Although our approach did not provide a large dataset for detail and comprehensive analysis of microbial community structure within the watershed, the analysis herein has provided us with very useful information on the role of certain environmental variables that shape community structure in low-flowing river. It is anticipated that as sequence technologies continue to advance, more data will be generated to understand the complexity and dynamics of microbial communities in low-flowing river.

Acknowledgements

We wish to express our appreciation to Pamela Watt, Greg Woodside, Nira Yamachika, Gary Hackney, and Bill Rice for their technical assistance and generous support during this study. We also thank Ms Phyllis Nash for statistical help. The study was made possible in part by the financial support of State Water Resources Control Board Nonpoint Source Pollution Control Programs Proposition 13. This research was also supported in part by the 206 Manure and Byproduct Utilization Project of the USDA-ARS. Mention of trademark or propriety products in this manuscript does not constitute a guarantee or warranty of the product by the USDA and does not imply its approval to the exclusion of other products that may also be suitable.

References

Abdo Z, Schüette UME, Bent SJ, Willians CJ, Forney LJ & Joyce P (2006) Statistical methods for characterizing diversity of microbial communities by analysis of terminal restriction fragment length polymorphisms of 16S rRNA genes. *Environ Microbiol* **8**: 929–938.

American Public Health Association (1995) *Standard Methods for the Examination of the Examination of Water and Wastewater*, 19th edn. American Public Health Association, Washington, DC.

Atlas R (1984) Use of microbial diversity measurements to assess environmental stress. *Current Perspectives in Microbial Ecology* (Klug MJ & Reddy CA, eds), pp. 540–545. American Society for Microbiology, Washington, DC.

Avaniss-Aghajani E, Jones K, Chapman D & Brunk C (1994) A molecular technique for identification of bacteria using small subunit ribosomal RNA sequences. *BioTechniques* **17**: 144–149.

Bernhard AE, Colbert D, McManus J & Field KG (2005) Microbial community dynamics based on 16S rRNA gene profiles in a pacific northwest estuary and its tributaries. *FEMS Microbiol Ecol* **52**: 115–128.

Blackwood CB, Marsh T, Kim SH & Paul EA (2003) Terminal restriction fragment length polymorphism data analysis for quantitative comparison of microbial communities. *Appl Environ Microbiol* **69**: 926–936.

Byappanahalli M, Fowler M, Shively D & Whitman R (2003) Ubiquity and persistence of *Escherichia coli* in a Midwestern coastal stream. *Appl Environ Microbiol* **69**: 4549–4555.

Byappanahalli M, Whitman RL, Shively DA, Sadowsky MJ & Ishii S (2006) Population structure, persistence, and seasonality of autochthonous *Escherichia coli* in temperate, coastal forest soil from a Great Lakes watershed. *Environ Microbiol* **8**: 504–513.

Crump BC & Hobbie JE (2005) Synchrony and seasonality in bacterioplankton communities of two temperate rivers. *Limnol Oceanogr* **50**: 1718–1729.

Crump BC, Kling GW, Bahr M & Hobbie JE (2003) Bacterioplankton community shifts in an Arctic lake correlate with seasonal changes in organic matter source. *Appl Environ Microbiol* **69**: 2253–2268.

Crump BC, Hopkinson CS, Sogin ML & Hobbie JE (2004) Microbial biogeography along an estuarine salinity gradient: combined influences of bacterial growth and residence time. *Appl Environ Microbiol* **70**: 1494–1505.

Dunbar J, Ticknor LO & Kuske CR (2000) Assessment of microbial diversity in four southwestern United States soils by 16S rRNA gene terminal restriction fragment analysis. *Appl Environ Microbiol* **66**: 2943–2950.

Dunbar J, Ticknor LO & Kuske CR (2001) Phylogenetic specificity and reproducibility and new method for analysis of terminal restriction fragment profiles of 16S rRNA genes from bacterial communities. *Appl Environ Microbiol* **67**: 190–197.

Fisher M, Graham JM & Graham LE (1998) Bacterial abundance and activity across sites within two northern Wisconsin *Sphagnum* bogs. *Microbial Ecol* **36**: 259–269.

Gentry RW, Layton AC, McKay LD, McCarthy JF, Williams DE, Koirala SR & Sayler GS (2006) *Escherichia coli* loading at or near baseflow in a mixed-use watershed. *J Environ Qual* **35**: 2244–2249.

Gentry RW, Layton AC, McKay LD, McCarthy JF, Williams DE, Koirala SR & Sayler GS (2007) Efficacy of Bacteroides measurements for reducing the statistical uncertainty associated with hydrologic flow and fecal loads in a mixed use watershed. *J Environ Qual* **36**: 1324–1330.

Gerdts G, Wichels A, Döpke H, Klings KW, Gunkel W & Schütt C (2004) 40-year long-term study of microbial parameters near Helgoland (German Bight, North Sea): historical view and future perspectives. *Helgol Mar Res* **58**: 230–242.

Hobbie JE, Daley RJ & Jasper S (1997) Use of nuclepore filters for counting bacteria by epifluorescence microscopy. *Appl Environ Microbiol* **33**: 1225–1228.

Hollister EB, Engledow ASM, Hammett AJ, Provin TL, Wilkinson HH & Gentry TJ (2010) Shifts in microbial community structure along an ecological gradient of hypersaline soils and sediments. *ISME J* **4**: 829–838.

© 2011 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved

Huber T, Faulkner G & Hugenholtz P (2004) Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* **20**: 2317–2319.

Izbicki JA, Pimentel MI & Leddy MB (2002) Microbial and Dissolved Organic Carbon characterization of stormflow in the Santa Ana River at Imperial Highway, Southern California, 1999–2002. U.S. Geological Survey, Scientific Investigation Report 2002; pp. 2004–5116.

Kent AD, Jones SE, Yannarell AC, Graham GM, Lauster GH, Kratz TK & Triplett EW (2004) Annual patterns in bacterioplankton community variability in a humic lake. *Microbial Ecol* **48**: 550–560.

LaMontagne MG, Schimel JP & Holden PA (2003) Comparison of subsurface and surface soil bacterial communities in California grassland as assessed by terminal restriction fragment length polymorphisms of PCR-amplified 16S rRNA genes. *Microbial Ecol* **46**: 216–227.

Langenheder S & Jurgens K (2001) Regulation of bacterial biomass and community structure by metazoan and protozoan predation. *Limnol Oceanogr* **46**: 121–134.

Liu WT, Marsh TL, Cheng H & Forney LJ (1997) Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. *Appl Environ Microbiol* **63**: 4516–4522.

Lozupone CA & Knight R (2007) Global patterns in bacterial diversity. *P Nat Acad Sci USA* **104**: 11436–11440.

Masin NM, Jezbera J, Nedoma J, Straskrabova V, Hejzlar J & Simek K (2003) Changes in bacterial community composition and microbial activities along the longitudinal axis of two canyon-shaped reservoirs with different inflow loading. *Hydrobiologia* **50**: 99–113.

McCune B & Mefford MJ (1999) *PC-ORD for Windows. Multivariate Analysis of Ecological Data. Version 4.01*. MJM, Gleneden Beach, OR.

Montgomery DC (2001) *Design and Analysis of Experiments*, 5th edn. John Wiley, New York, NY.

Murray AE, Preston CM, Massana R, Taylor LT, Blakis A, Wu K & DeLong EF (1998) Seasonal and spatial variability of bacterial and archaeal assemblages in the coastal waters near Anvers Island, Antarctica. *Appl Environ Microbiol* **64**: 2585–2595.

Neter J (1996) *Applied Linear Regression Models*, 3rd edn. Irwin, Chicago, IL.

Nogales B, Aguiló-Ferretjans MAM, Martín-Cardona C, Lalucat J & Bosch R (2007) Bacterial diversity, composition and dynamics in and around recreational coastal areas. *Environ Microbiol* **9**: 1913–1929.

Porter KG & Feig YS (1980) The use of DAPI for identifying and counting aquatic microflora. *Limnol Oceanogr* **25**: 943–948.

Rice WB (2005) Staff Report on Bacterial Indicator Total Maximum Daily Loads in the Middle Santa Ana River Watershed. Santa Ana Regional Water Quality Control Board, Santa Ana, CA.

Sapp MA, Wichels K, Wiltshire Hand & Gerdts G (2007) Bacterial community dynamics during the winter–spring transition in the North Sea. *FEMS Microbiol Ecol* **59**: 622–637.

SAS Institute (2005) *JMP User Guide*. Release 6. SAS Institute, Cary, NC.

Schafer H, Bernard L, Courties C *et al.* (2001) Microbial community dynamics in Mediterranean nutrient-enriched seawater mesocosms: changes in the genetic diversity of bacterial populations. *FEMS Microbiol Ecol* **34**: 243–253.

Schloss PD & Handelsman J (2006) Introducing treeclimber, a test to compare microbial community structures. *Appl Environ Microbiol* **72**: 2379–2384.

Schloss PD, Westcott SL, Ryabin T, *et al.* (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**: 7537–7541.

Sheneman L, Evans J & Foster JA (2006) Clearcut: a fast implementation of relaxed neighbor joining. *Bioinformatics* **22**: 2823–2824.

Smith CJ, Danilowicz BS, Clear AK, Costello FJ, Wilson B & Meijer WG (2005) T-Align, a web-based tool for comparison of multiple terminal restriction fragment length polymorphism profiles. *FEMS Microbiol Ecol* **54**: 375–380.

Sommer U, Gliwicz ZM, Lampert W & Duncan A (1986) The PEG-model of seasonal succession of planktonic events in fresh waters. *Arch Hydrobiol* **106**: 433–471.

Stepanauskas R, Moran MA, Bergamaschi BA & Hollibaugh JT (2003) Covariance of bacterioplankton composition and environmental variables in a temperate delta system. *Aquatic Microb Ecol* **31**: 85–98.

© 2011 Federation of European Microbiological Societies
Published by Blackwell Publishing Ltd. All rights reserved