Membrane and Water Treatment, *Vol. 10, No. 1 (2019) 029-038*
DOI: https://doi.org/10.12989/mwt.2019.10.1.029

Intra-event variability of bacterial composition in stormwater runoff from mixed land use and land cover catchment

Ma. Cristina A. Paule-Mercado[a], Imran Salim[b], Bum-Yeon Lee[a], Chang-Hee Lee* and Deokjin Jahng[c]

Department of Environmental Engineering and Energy, Myongji University, 116 Myongji-ro, Cheoin-gu, Yongin-si, Gyeonggi-do 17058, Republic of Korea

(Received May 7, 2018, Revised August 14, 2018, Accepted August 28, 2018)

Abstract. Microbial community and composition in stormwater runoff from mixed land use land cover (LULC) catchment with ongoing land development was diverse across the hydrological stage due different environmental parameters (hydrometeorological and physicochemical) and source of runoff. However, limited studies have been made for bacterial composition in this catchment. Therefore, this study aims to: (1) quantify the concentration of fecal indicator bacteria (FIB), stormwater quality and bacterial composition and structure according to hydrological stage; and (2) determine their correlation to environmental parameters. The 454 pyrosequencing was used to determine the bacterial community and composition; while Pearson's correlation was used to determine the correlation among parameters—FIB, stormwater quality, bacterial composition and structure—to environmental parameters. Results demonstrated that the initial and peak runoff has the highest concentration of FIB, stormwater quality and bacterial composition and structure. Proteobacteria, Bacteroidetes, Actinobacteria and Firmicutes were dominant bacteria identified in this catchment. Furthermore, the 20 most abundant genera were correlated with runoff duration, average rainfall intensity, runoff volume, runoff flow, temperature, pH, organic matter, nutrients, TSS and turbidity. An increase of FIB and stormwater quality concentration, diversity and richness of bacterial composition and structure in this study was possibly due to leakage from septic tanks, cesspools and latrines; feces of domestic and wild animals; and runoff from forest, destroyed septic system in land development site and urban LULC. Overall, this study will provide an evidence of hydrological stage impacts on the runoff microbiome environment and public health perspective.

Keywords: bacterial community; catchment; fecal indicator bacteria; land use and land cover; stormwater runoff

1. Introduction

Fecal indicator bacteria (FIB) contaminations are accumulated on the impervious or pervious cover within the watershed during dry days and it wash-off into the receiving water bodies during storm events. The *Escherichia coli* (EC), Enterococci (EN), fecal coliform (FC) and total coliform (TC) are commonly used to monitor the FIB contamination in stormwater runoff (Liao *et al*. 2015, Galfi *et al*. 2016, Ibekwe *et al*. 2016, Paule-Mercado *et al*. 2016, Staley *et al*. 2016, Bushon *et al*. 2018). The potential sources of these FIB in stormwater runoff include: (1) domestic and wildlife animal feces; (2) impervious and pervious cover; (3) land use and land cover (LULC) activities (biosolids or septage from agricultural and gardening operations; poor sanitation of solid waste collection; and land development, including removal of septic system, soil digging and soil transfer); and (4) infrastructure issues (aging sewer lines, illicit cross connection between stormwater runoff and sanitary sewer systems and industrial waste) (Liao *et al.* 2015, Galfi *et al.* 2016, Paule-Mercado *et al.* 2016, Staley *et al.* 2016, Bushon *et al.* 2018). An increase of FIB concentration in stormwater runoff has been leading causes for beach closure, directly linked to disease outbreaks and public health issues in watershed management (Liao *et al.* 2015, Stakey *et al.* 2016). Currently, the FIB studies were conducted in agricultural or urban watershed (Liao *et al.* 2015, Galfi *et al.* 2016, Huang *et al.* 2016, Paule-Mercado *et al.* 2016). These study have been made to determine the relationship between FIB and environmental parameters (hydrometeorological, LULC and physicochemical) and the influence of season or temperature on variability of FIB concentration. The occurrence of FIB and increase of its concentration in stormwater runoff does not always associate with the presence of pathogens and it does not give any indication of increase of fecal source contamination, respectively. Therefore, the inability to accurately estimate and identify the microbial community composition and structure in stormwater runoff will lead to inaccurate decision relating to public health.

In order to understand the microbial risk associated with stormwater runoff, comprehensive understanding on microbial community and composition and is needed. Also, studies investigating stormwater runoff microbiota using genetic markers or 16S rRNA gene-based microbial community surveys are useful tool to identify the occurrence of fecal contamination, bacterial community and

*Corresponding author, Ph.D., Professor
E-mail: changhee@mju.ac.kr
[a] Ph.D., Researcher
[b] Ph.D., Student
[c] Ph.D., Professor

ISSN: 2005-8624 (Print), 2092-7037 (Online)

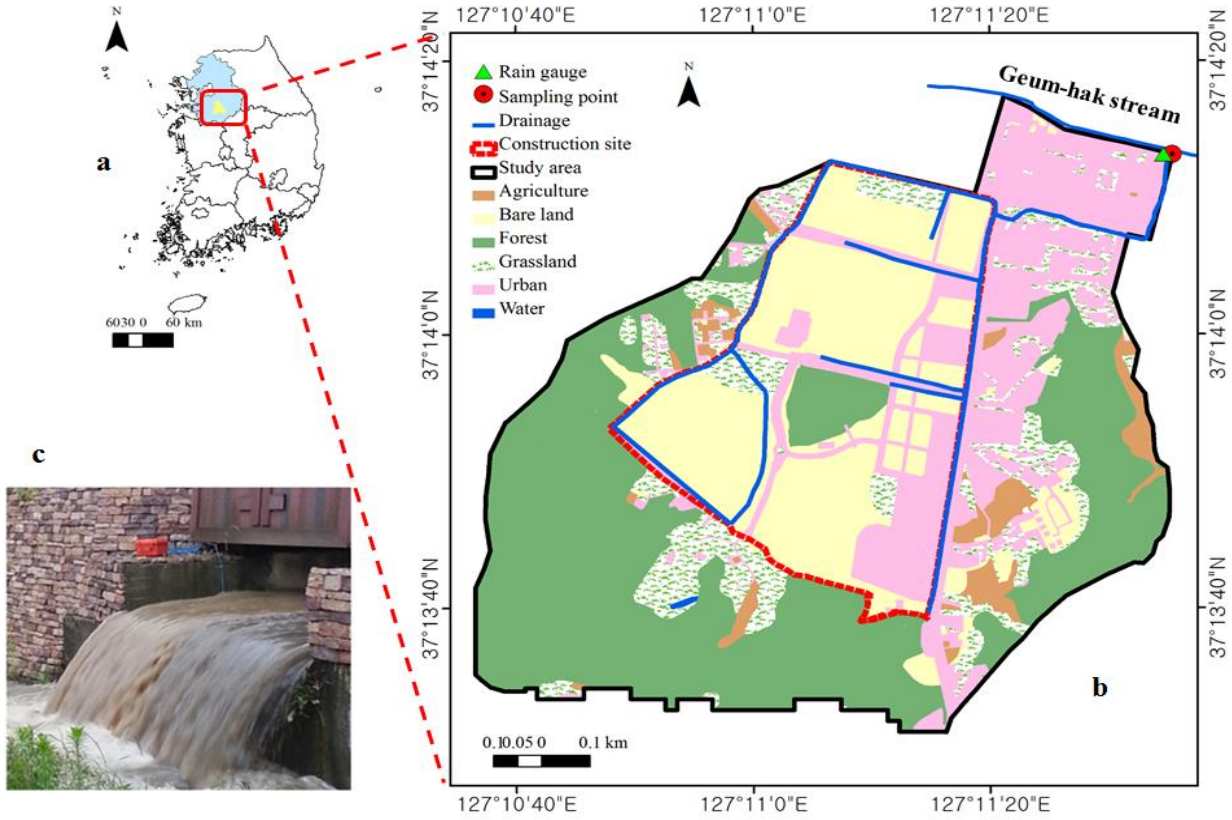


Fig. 1 Study area: (a) location; (b) LULC as of 2014; and (c) stormwater sampling site

composition and sources of bacterial contamination (Unno *et al.* 2010, Ibekwe *et al.* 2016, Ulrich *et al.* 2016). This will provide valuable information about the dominant microbial community and structure in stormwater runoff, enable deep analysis of inter- and intra-community correlations to environmental parameters and improve the bacterial source-specific identification in the receiving waterbodies (Ulrich *et al.* 2016). Therefore, the watershed managers will able to implement appropriate sustainable remediation strategies in order to improve the current water quality criteria. However, limited studies have been made for intra-event variability of bacterial community and composition in stormwater runoff from mixed LULC catchment with ongoing land development. Understanding the intra-event variability of bacterial composition and structure in stormwater runoff from this catchment will help to assess the overall public health risk and implementation of appropriate watershed management.

This paper aims to use the 16S rRNA genes in order to quantify the bacterial community composition in stormwater runoff collected from mixed LULC catchment with ongoing land development. Specifically, aims to: (1) quantify the intra-event variability of stormwater quality, FIB concentration, microbial community diversity, richness and composition; (2) identify the environmental parameters that explain the intra-event variability of stormwater quality, FIB, microbial community diversity, richness and composition; and (3) identify the relationship between environmental parameters and top twenty most abundant genera.

2. Materials and methods

2.1 Study area

The mixed LULC catchment is located within *Geum-Hak* watershed, Yongin City, Gyeonggi Province, South Korea (area: 1.45 km^2; location: between 37°13’29.85”N-37°14’12.9”N and 127°10’16.97”E-127°11’29.78”E) (Figs. 1(a)-(b)). Stormwater runoff from construction site, forest and urban LULC are directly discharge into the outfall (Fig. 1c) of the catchment and eventually into *Geum-hak* stream (one of the tributaries of the Paldang reservoir, which serve as the main source of drinking water of Seoul City and nearby provinces) (Paule *et al.* 2014). Climatically, this catchment has subtropical monsoon with four distinct seasons: spring (March to May; -1.5 to 23.4°C); summer (June to August; 16.3 to 31.9°C); autumn (September to November; -2.3 to 27.4°C); and winter (December to February; -13.3 to 7.6°C). Between 2003 and 2014, the mean annual rainfall was 1,457.9 mm (85% of the rainfall occurred during summer). The upstream area has ongoing LULC development (from agricultural and forest into residential and commercial complexes), whereas the downstream area is about 95% urban. As of 2014, the major LULC was bare land (53.82%), followed by forest (22.62%), urban (12.04%), grassland (7.29%), agriculture (3.0%) and water (1.23%) (Fig. 1b). The monitoring site has a separate sewer system (>20 years) and it is sufficient for exfiltration and infiltration of stormwater system. The possible source of bacteria in this catchment was feaces from domestic and wild animals, cesspools and latrines,

runoff from forest, destroyed septic system in land development site and urban LULC.

2.2 Sampling strategy, collection and analysis

Between May 2013 and October 2014, a total of 16 storm events and grab stormwater samples (n=80) at the outlet of the catchment were monitored and analysed, respectively. Monitoring was initiated when there is ≥3 days of ADD, 4 mm of total rainfall and 6 h total rainfall duration to ensure that there is enough runoff flow, pollutant build-up and wash-off. Stormwater samples were collected according to time and flow weighted sampling, which allow to evaluate the occurrence of first-flush and intra-event pollutant concentration variations. In all storm events, sampling was started when runoff was generated or a flow above baseline conditions was observed during a rainfall event. If sampling site has base flow, single sample was taken before the rain to exclude background pollution impact. Small sampling intervals (5 min to 15 min) were considered during the initial storm period (first two hours) and during peak flow conditions, where large sampling intervals (1 to 2 hr) were considered for remaining storm duration. Due to limited resources of this study, however, it was not possible to characterize all collected samples for the FIB concentration and bacterial community and composition. Therefore, the interval and number of stormwater samples were selected in the laboratory according to the hydrological stage (initial, peak and final), which is time, runoff flow and turbidity base, as shown in Table 1. The initial sample (1 sample selected) was collected between 0 and 30 min interval. Peak samples were selected 2 to 5 times: 2 to 3 samples were collected between 120 and 240 min; and 2 to 3 samples were collected between 300 and 360 min. This is because during monitoring it was observed that the initial peak was influenced by the surrounding urban LULC; while the runoff from land development and other pervious cover resulted in later peak. The final sample (1 sample selected) was collected ≥ 360 min interval or until the runoff became visually clear (≤ 30NTU) with decrease in flow rate. This strategy was made to ensure that the selected samples analysed the influence of first flush, peak runoff and different LULC (e.g. urban, land development, etc.).

Continuous flow (PCM F NIVUS, Germany; 1 min interval) and rainfall data (HB-3207-09, Casella, UK; 1 min interval) were measured in the monitoring site (Fig. 1c). The ADD, average rainfall intensity and total rainfall were acquired using the Suwon rain gauge station (about 19 km away from the outfall) of the Korean Meteorological Administration (http://web.kma.go.kr/eng/index.jsp).

Stormwater samples were collected (4 L sterile polyethylene bottle), stored (4°C), transported into the laboratory and analyzed within 6 hr of collection. Conductivity, pH and stormwater temperature were measured *in situ* using the calibrated multiprobe (Horiba U-50 Multi-Probe, Japan). Turbidity was determined using calibrated Hach 2100AN Portable Turbidimeter (Loveland, CO). The 5-day biochemical oxygen demand (BOD_5), chemical oxygen demand (COD), chloride (Cl), total nitrogen (TN), total phosphorus (TP) and total suspended solids (TSS) were analyzed according to standard methods for examining water and wastewater (APHA, 2005, 2009). Enumeration of FIB (*Escherichia coli* (EC); Enterococci (EN); fecal coliform (FC); and total coliform (TC)) concentrations were determined using the American (APHA, 2005 and 2009) and modified Korean (Korea Ministry of Environment, 2011) standards for the examination of water. FIB measured via the Colilert and Enterolert Defined Substrate Technology kits (IDEXX Laboratories, Inc., Westbrook, ME, USA) and were enumerated as most probable number (MPN) per 100 mL of water by the Quanti-Tray Method (IDDEXX Laboratories, Maine, USA) (APHA, 2009). Paule-Mercado *et al*. (2016) described the detailed FIB enumerations used in this study.

Table 1 Sampling interval and number of samples collected

Hydrological Stage	Sampling interval (min)	Average runoff flow (log_{10} m^3/hr)	Average turbidity (log_{10} NTU)	Number of samples	Total number of samples collected
Initial	0 to120	2.27	1.03	1	16
Peak	120 to 240 and 300 to 360	3.92	2.32	2 to 5	48
Final	>360	2.06	0.85	1	16
Sum					80

2.3 DNA extraction, purification and pyrosequencing

Each selected stormwater samples (1 L) was filtered (0.45-µm-pore-size mixed cellulose ester filter; 47 mm diameter) for DNA collection. Filters were folded, placed in 2 mL tubes and stored at -80°C (< 1 week), until DNA extraction was performed. To extract the DNA from frozen filters, the filters were removed from tubes and -80°C and crushed into fine powder (< 2 mm diameter) with the use of sterile spatulas. The DNA was extracted using a FastDNA Spin Kit for Soil (MP Biomedicals, Solon, OH) according to manufacturer's guidelines. The DNA quality was checked by electrophoresis on 1% agarose gels.

The extracted DNA was amplified using primers targeting the V1 to V3 regions of the 16S r RNA. As primers, the bacterial primers of 27F (5'-GAGTTTGATCMTGGCTCAG-3') and 518R (5'-ATTACCGCGGCTGCTGG3') were used. The reverse primer was attached in adapter 1 (5'-CCATCTCATCCCTGCGTGTCTCCGAC-3'), core nucleotides (TCAG), barcode (7-11 nucleotides) and a linker (AC). The forward primer was attached by an adapter 2 (5'-CCTATCCCCTGTGTGCCTTGGCAGTC-3'), core nucleotides and a linker. The Bifidobacterium, however, cannot amplify with the universal primers, specifically the B16S-F (CCTATCCCCTGTGTGCCTTGGCAGTC-TCAG-AC-GAGTTTGATCMTGGCTCAG). Therefore, mixtures of B16S-F and Bif16S-F (CCTATCCCCTGTGTGCCTTGGCAGTC-TCAG-AC-GG GTTCGATTCTGGCTCAG), with 9:1 ratio, were used for community analysis of all stormwater samples. These mixed primers were used with all other barcoded reverse primers. The PCRs were carried out under the following

conditions: initial denaturation at 94°C for 5 min; followed by 30 cycles of 94°C for 30 s (denaturation), 55°C for 30 s (annealing) and 72°C for 30 s (extension), with a final extension at 72°C for 7 min. Amplified products were confirmed by 1.2% agarose gel electrophoresis and purified with MO BIO PowderClean DNA cleanup kit (MO BIO Laboratories Inc., Carlsbad, CA) according to manufacturer's guidelines. The DNA concentrations were quantified using a NanoDrop ND-1000 spectrometer (Thermo Scientific, Wilmington DE). The DNA sequencing was performed by ChunLab, Inc (Seoul, South Korea), with a Roche/454 GS FLX Titanium platform, according to the manufacturer's guidelines.

The sequencing reads separated by barcode, 2 bp of linker or primers were trimmed. Sequencing reads with low quality (average quality score < 25 or read length < 300) were filtered. The 16S rRNA amplicons have no significant match in a Basic Local Alignment Search Tool (BLAST) search (exceptional value of > e-^{5}) against the Gene Bank database were also removed for further analyses. Noise and chimeras sequences were omitted from the dataset using the UCHIME program. For taxonomic assignment of each pyrosequencing reading, EzTaxon-e database (http://extaxon-e.ezbiocloud.net/) was used. In this study, each reads were taxonomically assigned according to following criteria: genus (97% > x ≥ 94%), family (94% > x ≥ 90%) and phylum (80% > x ≥ 75%). The read is assigned as an unclassified group if the similarity was below the cutoff ranks. The CLcommunity software provided by the ChunLab (http://www.chunlab.com/) was used to assess the species diversity and richness. For microbial diversity, Operational taxonomic units (OTUs; identity cut-off of 97%) were defined by using the CD-HIT program. The species richness/evenness was estimated using ACE, Chao1, Shannon and Simpson indices by using the MOTHUR program.

2.4 Statistical and pyrosequencing data analysis

All parameters (FIB, environmental and abundance (%) of each taxon) were $\log_{10}$-transformed prior to analysis. The following environmental parameters were used in this study: hydrometeorological—runoff flow, average rainfall intensity, runoff volume and temperature; and physicochemical—pH, conductivity, Cl, BOD_5, COD, TN, TP, TSS and turbidity. Pearson's correlation was used to identify the significant correlation between: (1) physicochemical and hydrometeorological; (2) FIB and physicochemical; (3) among FIB parameters; (4) microbial diversity and environmental parameters; and (5) microbial community structure (20 most abundant genera) and environmental parameters. All statistical analyses were conducted in SigmaPlot 12.3 (Systat Software, Inc., San Jose, CA) with $p < 0.5$ significance level.

3. Results and discussion

3.1 Stormwater quality

Stormwater quality concentrations were varied according to hydrological stage (Fig. 2). The concentrations of conductivity, Cl, BOD_5, COD, TN and TP were higher during initial samples; temperature and concentrations of TSS and turbidity during peak; and only pH was higher concentration in final sampling. The initial concentration of Cl, BOD_5, COD, TN and TP have positive significant correlation with FIB ($r > 0.851$, $p < 0.05$), stormwater runoff ($r > 0.878$, $p < 0.05$), average rainfall intensity ($r > 0.833$, $p < 0.05$) and runoff volume ($r > 0.821$, $p < 0.05$). The peak concentration of BOD_5, COD, TN, TP, TSS and turbidity have positive significant correlation with FIB ($r > 0.872$, $p < 0.05$), stormwater runoff ($r > 0.845$, $p < 0.05$), rainfall intensity ($r > 0.828$, $p < 0.05$) and runoff volume ($r > 0.837$, $p < 0.05$). Weak significant correlation was observed between the final concentrations of pollutants with FIB ($r > 0.421$, $p < 0.05$), stormwater runoff ($r > 0.681$, $p > 0.05$), rainfall intensity ($r > 0.384$, $p < 0.05$) and runoff volume ($r > 0.551$, $p > 0.05$). These results suggest that the catchment has high organic matter, nutrients, solids and bacterial concentration at the initial and peak than final of events. Also, the initial and peak concentration of pollutants in stormwater runoff were influenced by illicit cross connections between stormwater and sanitary sewer system, farming practices (e.g. application of fertilizer before the storm events), destroyed septic system due to construction activities, land alteration (e.g. soil digging and soil transfer), forest and impervious cover (Galfi *et al.* 2016, Huang *et al.* 2016, Ibekwe *et al.* 2016, Paule-Mercado *et al.* 2016, Bushon *et al.* 2018). Moreover, majority of the collected pollutant concentrations were above the recreational water quality criteria (USEPA, 2014, MOE 2014). Therefore, the stormwater runoff in the catchment may contribute to the diffuse pollution of the receiving streams.

3.2 FIB concentrations

The FIB concentrations varied according to hydrological stage of storm events (Fig. 3). The concentration of TC ranged from 3.86 (final) to 7.17 (peak) $\log_{10}$ MPN/100 mL; FC ranged from 3.57 (final) to 6.99 (peak) $\log_{10}$ MPN/100 mL; EC ranged from 2.37 (final) to 6.94 (peak) $\log_{10}$ MPN/100 mL; and EN ranged from 2.41 (initial) to 6.73 (peak) $\log_{10}$ MPN/100 mL. Overall, there was small significance difference in FIB concentrations between initial (3.4 - 6.95 $\log_{10}$ MPN/100 mL) and peak (4.32 - 6.99 $\log_{10}$ MPN/100 mL) concentration (p=0.09), while the final concentration showed lower concentration (2.36 - 6.95 $\log_{10}$ MPN/100 mL).

Result on intra-event variability of FIB concentrations were expected due to first flush and increase of sewage release during monitoring at initial and peak runoff. Leakage from septic tanks, cesspools and latrines; feces of domestic and wild animals; runoff from forest, destroyed septic system in land development site and urban areas may contribute to elevate the intra-event contamination of FIB in stormwater runoff. Also the other potential factors, rainfall duration ($r > 0.852$, $p < 0.05$) and ADD ($r > 0.848$, $p < 0.05$) might be expected to have a high influence on FIB. The FIB concentration during peak runoff was higher than those during initial runoff although their ADD was almost similar. This could be explained by environmental conditions such as stormwater temperature. Since

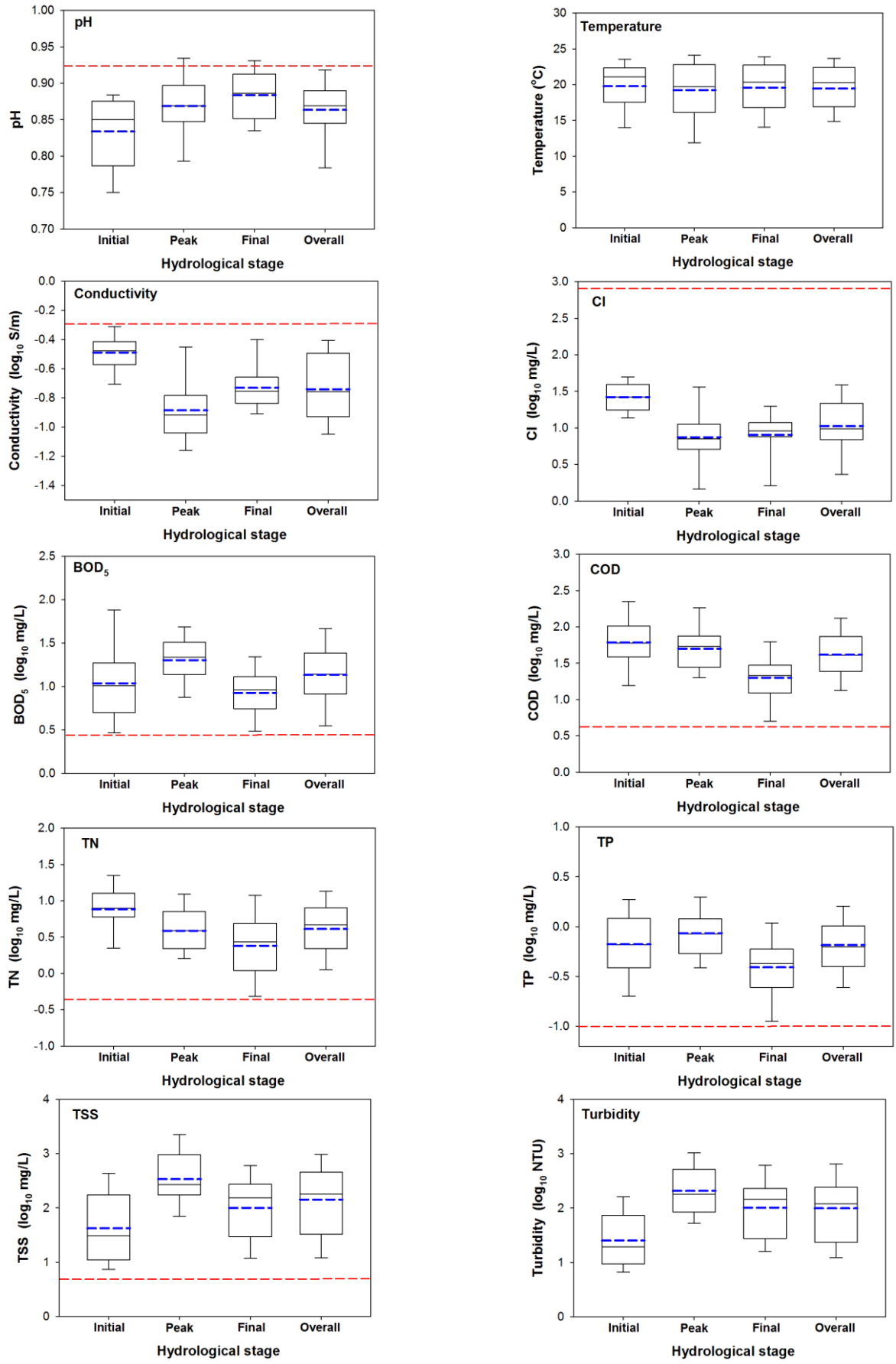


Fig. 2 Intra-event concentrations of FIB (log_{10} mg/L). The blue and red dashed lines represent the mean concentration and recreational water quality standard (USEPA 2012, MOE 2014) for FIB, respectively

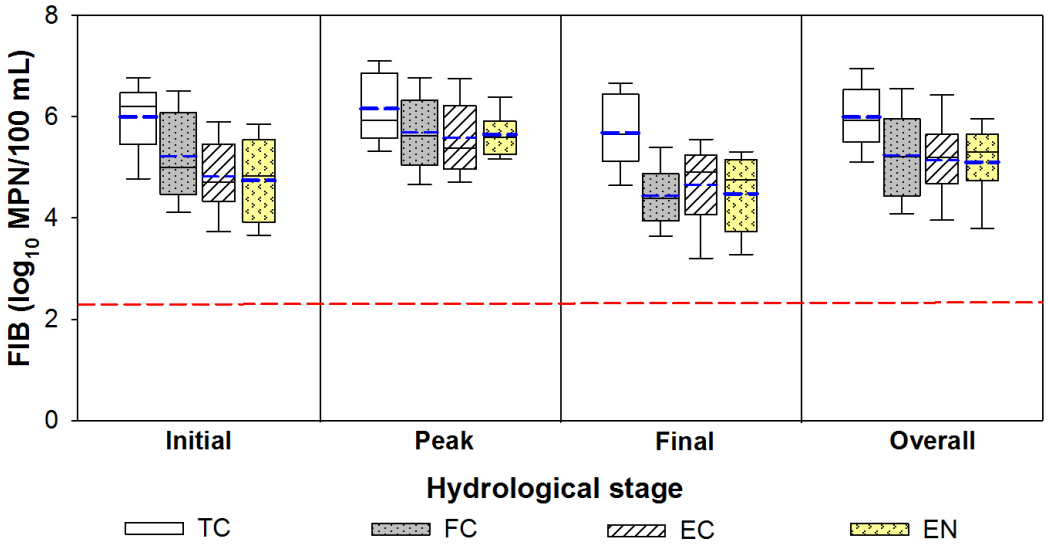


Fig. 3 Intra-event concentrations of FIB (log_{10} mg/L). The blue and red dashed lines represent the mean concentration and recreational water quality standard (USEPA 2012, MOE 2014) for FIB, respectively

temperature has impacts on both biological and physicochemical reactions (Huang *et al.* 2016). Moreover, according to the USEPA (2012) and MOE (2014) guidelines for recreational water, the concentration of FIB should be below 230 (2.361 log_{10}) MPN/100 mL in any of the samples tested. Based on these standard guidelines, all stormwater samples collected in this study were considered unsuitable for recreational purposes. Thus, the runoff in this catchment may contribute to the diffuse microbial pollution of the receiving water bodies.

3.3 Microbial community diversity and richness

Table 2 shows the summary of the number of sequence (Nseq) tags, OTUs, Ace, Chao1, Shannon and Simpson. Overall, a total of 335,123 16S rRNA Nseq were generated through 454 pyrosequencing, with an average length of about 210 bp. The 454 sequence libraries (total) ranged from 92,598 (final) to 126,308 (peak) sequences and contained between 18,159 (final) to 22,136 (peak) OTUs. All collected data were normalized into smallest Nseq and reanalyzed to determine the normal distribution of variances.

The Ace values suggest that the diversity varied according to hydrological stage, with the lowest and highest diversity associated with final (53,923) and peak (68,970) runoff, respectively. The total Chao1 richness was higher sequences during peak (46,920.01), followed by initial (39,926.89) and final (36,239.49) runoff. Shannon's diversity index values showed that the peak (4.568 to 6.871) had the highest microbial diversity, followed by initial (4.312-6.408) and final (4.020-6.797) runoff and there were significant differences ($r > 0.862$, $p < 0.05$) among the hydrological stages. In terms of Simpson's diversity index, the peak (0.251) had the highest diversity, followed by initial (0.273) and final (0.303) runoff. The diversity values increased significantly ($r > 0.812$, $p < 0.05$) across the hydrological stages. OTUs, Shannon and Simpson represent

Table 2 Descriptive statistics of sequence library, OTUs, diversity and richness at 97% level

Parameters		Hydrological stage			
		Initial	Peak	Final	Overall
Nseq*	Min	3,561	3,614	2,915	2,915
	Max	12,237	15,318	8,531	15,318
	Average	7,264	8,019	5,787	7,023
	Sum	116,217	126,308	92,598	335,123
OTUs**	Min	767	890	696	696
	Max	2,582	2,704	2,338	2,704
	Average	1,242	1,384	1,135	1,254
	Sum	19,870	22,136	18,159	60,165
Ace	Min	1,733	2,441	1,631	1,631
	Max	8,502	8,899	8,656	8,899
	Average	3,530	4,311	3,370	3,737
	Sum	56,478	68,970	53,923	179,371
Chao1	Min	1,350.41	1,678.38	1,225.12	1,225.12
	Max	5,689.51	5,719.50	5,389.50	5,719.50
	Average	2,495.43	2,932.50	2,264.97	2,564.30
	Sum	39,926.89	46,920.01	36,239.49	123,086.39
Shannon	Min	4.312	4.568	4.020	4.020
	Max	6.408	6.871	6.797	6.871
	Average	5.592	5.899	5.512	5.668
	Sum	89.462	94.365	88.186	272.013
Simpson	Min	0.004	0.003	0.005	0.003
	Max	0.058	0.046	0.062	0.062
	Average	0.017	0.016	0.019	0.017
	Sum	0.273	0.251	0.303	0.827

*Nseq = number of sequence tags; **OTUs = operational taxonomic units

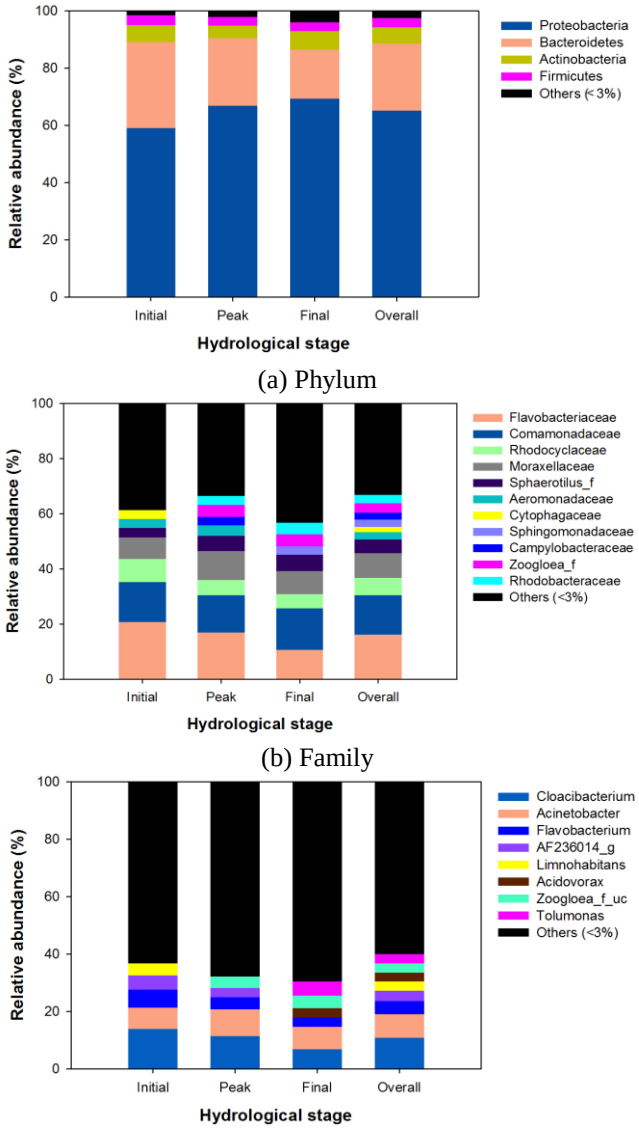
the microbial diversity of the community. The higher the OTUs and Shannon value, the more diverse the microbial community is, and the lower Simpson value indicates an abundant communities. However, Ace and Chao1 values are the indices of the microbial richness. The higher Nseq, Ace

and Chao1 values shows a more diverse microbial structure. So, the microbial diversity in this catchment was in the following order: peak > initial > final based on the OTUs, Shannon and Simpson values. Community richness were in the order of peak > initial > final according to Ace and Chao 1. The main reason for peak sample had the most diverse and abundant microbial communities compared to the other hydrological stage could be that this stage has 90% of the combined runoff from surrounding urban LULC, forest and land development area. The runoff volume during this stage was more than 3.82 $\log_{10}$ m^3, resulting in microbial and pollutants wash-off and less exposure to solar radiation. Also, peak runoff has temperature ranged from 11.2°C to 23.7°C and abundant to nutrients, TSS and organic matter, which suitable for microbial environment.

3.4 Microbial community structure

Bacterial community profiles were generated for the mixed LULC catchment. Fifty-one phyla were represented among the sequenced samples, with the most abundant (average abundance > 3%) taxa included Proteobacteria, (59.02%, initial; 66.85%, peak; 69.29%, final; 65.06%, overall), Bacteroidetes (30.05%, initial; 23.65%, peak; 17.13%, final; 23.61%, overall), Actinobacteria (6.07%, initial; 4.44%, peak; 6.55%, final; 5.69%, overall) and Firmicutes (3.41%, initial; 3.01%, peak; 3.10%, final; 3.16%, overall) (Fig. 4a). These phyla have been previously identified within stormwater runoff and their abundance has been associated with the presence of untreated sewage, human or animal fecal microorganisms and contribution of sediment erosion (Leung *et al.* 2005, Shanks *et al.* 2013, Mcllelan *et al.* 2015).

The dominant families (average abundance > 3%) were different across the hydrological stage (Fig. 4b). Among the Proteobacteria, the initial sample has the most abundant family of Comamonadaceae (14.50%), Rhodocyclaceae (8.26%), Moraxellaceae (7.87%), Sphaerotilus)_f (3.40%) and Aeromonadaceae (3.31%). Within Bacteroidetes, Flavobacteriaceae (20.79%) and Cytophagaceae (3.23%) were the major family group. In peak sample, the major family groups were Comamonadaceae (13.43%), Moraxellaceae (10.62%), Rhodocyclaceae (5.53%), Sphaerotilus_f (5.51%), Zoogloea_f (4.28%), Aeromonadaceae (3.84%), Rhodobacteraceae (3.32%) and Campylobacteraceae (3.04%). Flavobacteriaceae (16.98%) constituted the major family group among Bacteroidetes. In final runoff, Comamonadaceae (14.95%), Moraxellaceae (8.35%), Sphaerotilus_f (5.97%), Rhodocyclaceae (5.19%), Zoogloea_f (4.41%), Rhodobacteraceae (4.17%) and Sphingomonadaceae (3.14%) were the major family group in Proteobacteria. Flavobacteriaceae (10.67%) represented the most of Bacteroidetes in the final runoff. Results shows that the increase of type and abundance of microbial family level were mirrored by increased the sampling period. For instance, Sphingomonadaceae, Campylobacteraceae, Zoogloea_f and Rhodobacteraceae not found in initial runoff but present either in peak or final runoff. The main reason was due to difference in environmental factors such as source of runoff, pH, temperature, salinity, organic matter, nutrients, TSS and turbidity during hydrological



(a) Phylum

(b) Family

(c) Genera

Fig. 4 Average phylogenic distribution of each microbial community according to hydrological stage

stage. Microbial communities at initial runoff influenced by the untreated sewage or households waste and surrounding urban LULC or impervious cover, while microbial communities at peak or final runoff may be influenced by the runoff from forest, land development, destroyed septic system in construction site and soil alteration (soil digging and soil transfer from forest to bare land). Since the runoff from these areas (generally pervious cover) required high rainfall intensity and runoff duration before the pollutants will wash-off.

Overall, eight genera (average abundance > 3%) including Cloacibacterium (10.77%), Acinetobacter (8.18%), Flavobacterium (4.69%), AF236014_g (3.55%), Tolumonas (3.37%), Limnohabitans (3.24%), Zoogloea_f_uc (3.20%) and Acidovorax (3.01%) dominated (average abundance > 3%) the bacterial communities of stormwater runoff according to hydrological stage (Fig. 4c). At the initial runoff, the Acinetobacter (7.51%), AF236014_g (4.93%) and

Table 3 Pearson correlations between environmental parameters and 20 most abundant genera from mixed LULC catchment

Genera	Rank	ADD	Runoff duration	Average rainfall intensity	Runoff volume	Flow	Temp	pH	Cl	BOD_5	COD	TN	TP	TSS	Turbidity
Acidovorax	7	**0.421**	**0.573**	-0.405	**0.521**	**0.578**	**0.801**	**0.536**	0.441	**0.511**	**0.514**	**0.578**	**0.564**	**0.515**	**0.542**
Acinetobacter	2	**0.401**	**0.509**	**0.603**	**0.586**	**0.602**	**0.605**	**0.502**	**0.473**	**0.747**	**0.536**	**0.693**	**0.511**	**0.668**	**0.753**
Aeromonas	20	**0.423**	**0.524**	-0.475	**0.528**	**0.512**	**0.802**	**0.517**	**0.446**	**0.852**	**0.864**	**0.893**	**0.825**	**0.537**	**0.522**
AF236014_g	4	**0.425**	**0.598**	**0.515**	**0.545**	**0.601**	**0.821**	**0.525**	**0.447**	**0.685**	**0.626**	**0.713**	**0.583**	**0.723**	**0.792**
Aquabacterium	19	**0.435**	**0.506**	**0.602**	**0.513**	**0.502**	**0.879**	**0.514**	**0.411**	**0.522**	**0.571**	**0.508**	**0.554**	**0.532**	**0.616**
Arcobacter	11	**0.489**	**0.557**	-0.222	**0.504**	**0.554**	**0.817**	**0.537**	0.777	**0.851**	**0.828**	**0.527**	**0.542**	**0.554**	**0.587**
Cloacibacterium	1	**0.415**	**0.511**	**0.558**	**0.513**	**0.513**	**0.683**	**0.537**	**0.453**	**0.761**	**0.724**	**0.606**	**0.598**	**0.688**	**0.736**
Dechloromonas	14	0.504	**0.541**	**0.607**	**0.546**	**0.607**	**0.517**	**0.525**	**0.434**	**0.805**	**0.822**	**0.579**	**0.594**	**0.593**	**0.526**
Flavobacterium	3	**0.433**	**0.555**	0.466	**0.601**	**0.516**	**0.542**	**0.547**	**0.456**	**0.718**	**0.654**	**0.505**	**0.623**	**0.548**	**0.587**
Hydrogenophaga	10	0.536	**0.526**	-0.014	**0.542**	**0.582**	**0.802**	0.431	**0.478**	**0.533**	**0.584**	**0.507**	**0.715**	**0.599**	**0.578**
JN679217_g	18	0.547	**0.537**	0.435	**0.531**	**0.567**	**0.804**	**0.509**	**0.474**	**0.597**	**0.578**	**0.525**	**0.599**	**0.521**	**0.543**
Limnohabitans	8	**0.411**	**0.601**	**0.602**	**0.604**	**0.545**	**0.863**	**0.506**	**0.483**	**0.522**	**0.567**	**0.513**	**0.593**	**0.511**	**0.572**
Mycobacterium	16	**0.492**	**0.603**	-0.422	**0.608**	**0.502**	**0.864**	**0.503**	0.569	**0.507**	**0.519**	**0.564**	**0.532**	**0.829**	**0.809**
Novosphingobium	13	**0.483**	**0.579**	0.205	**0.517**	**0.584**	**0.533**	**0.588**	0.577	**0.831**	**0.588**	**0.529**	**0.741**	**0.568**	**0.565**
Polynucleobacter	15	**0.435**	**0.602**	-0.499	**0.518**	**0.534**	**0.759**	**0.577**	0.795	**0.522**	**0.519**	**0.726**	**0.573**	**0.563**	**0.577**
Pseudarcicella	17	**0.445**	**0.518**	0.411	**0.561**	**0.509**	**0.884**	**0.526**	**0.458**	**0.653**	**0.657**	**0.516**	**0.508**	**0.866**	**0.749**
Rhodobacter	12	**0.478**	**0.545**	-0.212	**0.521**	**0.591**	**0.523**	**0.519**	**0.472**	**0.827**	**0.787**	**0.588**	**0.563**	**0.573**	**0.574**
Sphaerotilus	9	**0.425**	**0.592**	**0.598**	**0.593**	**0.547**	**0.871**	**0.525**	**0.483**	**0.527**	**0.564**	**0.752**	**0.568**	**0.584**	**0.534**
Tolumonas	5	**0.441**	**0.606**	**0.523**	**0.562**	**0.509**	**0.826**	**0.575**	**0.415**	**0.545**	**0.618**	**0.558**	**0.718**	**0.562**	**0.564**
Zoogloea	6	**0.474**	**0.528**	**0.575**	**0.506**	**0.512**	**0.875**	**0.568**	**0.421**	**0.517**	**0.565**	**0.891**	**0.514**	**0.725**	**0.719**
Minimum		7	**0.421**	**0.573**	-0.405	**0.521**	**0.578**	**0.801**	**0.536**	0.441	**0.511**	**0.514**	**0.578**	**0.564**	**0.515**
Maximum		2	**0.401**	**0.509**	**0.603**	**0.586**	**0.602**	**0.605**	**0.502**	**0.473**	**0.747**	**0.536**	**0.693**	**0.511**	**0.668**

All environmental parameter and genera were $\log_{10}$ transformed prior to statistical analyses to meet the normality requirement; Rank, rank of bacterial composition at genus level; minimum and maximum significant values; all bold values are significant at $p < 0.05$

Limnohabitans (4.06%) were the most dominant genera within the Proteobacteria; while Cloacibacterium (13.83%) and Flavobacterium (6.33%) were the only dominant genera for Bacteroidetes. For peak runoff, the Acinetobacter (9.31%), Zoogloea_f_uc (3.99%) and AF236014_g (3.08%) were the dominant genera within the Proteobacteria. However, Cloacibacterium (11.32%) and Flavobacterium (4.36%) were the dominant genera within the Bacteroidetes. For final runoff, the Acinetobacter (7.68%), Acidovorax (3.23%), Tolumonas (4.93%) and Zoogloea_f_uc (4.31%) were the major genera of Proteobacteria; while Cloacibacterium (6.89%) and Flavobacterium (3.38%) for Bacteroidetes. Similar to higher taxonomic classification (phylum and family), the microbial communities at genus level were different according hydrological stage. The environmental factors contributed to the occurrence, spread and persistence of various bacterial communities at genus level in the stormwater runoff. For instance, the Acidovorax, Acinetobacter, Cloacibacterium, Flavobacterium, Limnohabitans, Tolumonas and Zoogloea are commonly associated with urban runoff, wastewater treatment plants, sanitary sewage (Shanks *et al.* 2013, Mcllelan *et al.* 2015, Brinkmeyer 2016, Eren *et al.* 2016).

3.5 Microbial community and environmental parameters

Table 3 shows the correlation between top 20 most abundant genera and environmental parameters. Overall, the bacterial communities and environmental parameters have weak positive correlation with ADD ($r \leq 0.492$, $p < 0.05$) and Cl ($r \leq 0.483$, $p < 0.05$). Therefore, the bacterial communities do not require more than three days for buildup and it will wash-off during the stormwater runoff and it will reach the catchment outlet. Also, the chloride concentration in stormwater runoff was not sufficient to reduce the bacterial communities.

Runoff duration ($r \leq 0.606$, $p < 0.05$), average rainfall intensity ($r \leq 0.607$, $p < 0.05$), runoff volume ($r \leq 0.608$, $p < 0.05$) and runoff flow ($r \leq 0.607$, $p < 0.05$) have moderate positive correlation with bacterial communities. This result together with storm wash-off off of bacterial communities is a key process in controlling the storm event concentration of bacteria, which is not surprising in a small with mixed LULC catchment. Also, the positive association between these variables suggested that light storm events (e.g. 4 to 10 mm) with long runoff duration can introduce high bacterial communities from human sources than the

storm events with high average rainfall intensity. This result could potentially occur if the source of bacterial contamination is a wastewater treatment plant and illicit or leaking connections of aging sewer lines into stormwater runoff, as compared to bacterial contamination from surface sources (e.g. animals or biosolids) (Liao *et al.* 2015, Bushon *et al.* 2018).

Stormwater temperature ($r \leq 0.884$, $p < 0.05$) and pH ($r \leq 0.588$, $p < 0.05$) have strong and moderate positive correlation to bacterial communities, respectively. This result suggests that the temperature (9.4°C to 24.5°C) and pH (5.25 to 9.12) condition in the stormwater collected influenced the life cycle of bacterial communities, specifically the fate, growth, reproduction and survival rates (Cabral *et al.* 2018).

Organic matter ($r \leq 0.864$, $p < 0.05$), nutrients ($r \leq 0.893$, $p < 0.05$), TSS ($r \leq 0.866$, $p < 0.05$) and turbidity ($r \leq 0.803$, $p < 0.05$) has a strong positive correlation to bacterial communities. This correlation is not surprising, because the urban runoff has the potential to pollute the water in different ways such as illicit sewer connections or leaks in sanitary sewer and wastewater treatment plant as this may contain pollutants such as FIB, nutrients, organic matter, TSS and other bacteria (Ibekwe *et al.* 2016, Paule-Mercado *et al.* 2016). Also, the runoff from construction site, forest and land development area influenced the association between these parameters and bacterial communities.

4. Conclusions

This study emphasizes the complexity of the environmental parameters (hydrometeorological and physicochemical) that can influence the intra-event variability of FIB and microbial community in stormwater samples collected over 2-year period from mixed LULC catchment within *Geum-Hak* stream, in Yongin City, Gyeonggi Province, South Korea. The following conclusions can be made:

- Generally, the initial and peak runoff have the highest concentration and with strong positive correlation to FIB and environmental parameters.
- Using the pyrosequencing, initial and peak runoff have higher Proteobacteria, Bacteroidetes, Actinobacteria and Firmicutes major group identified.
- An increase of FIB, environmental parameters and microbial communities was possibly due to leakage from septic tanks, cesspools and latrines; feces of domestic and wild animals; runoff from forest, destroyed septic system in land development site and urban areas.
- The identified specific group or level (phylum, family and genus) in stormwater samples can be used as a basis for assessing the relative importance on the increase microbial communities and possible source of microbial communities (through literature review).
- The top twenty most abundant genera identified in this study was generally correlated with runoff duration, average rainfall intensity, runoff volume, runoff flow, temperature, pH, organic matter, nutrients, TSS and turbidity.
- Moreover, comparison with other catchment (e.g. urban catchment), association or correlation of LULC and particle size distribution and using high-throughput molecular methods (e.g. microbial source tracking) will advance the understanding of microbial communities in stormwater runoff and it will improve the preventive measures for public health.

Acknowledgments

This research was supported by the Ministry of Oceans and Fisheries, Korea (Development of Integrated Estuarine Management System) and Korea Environmental Technology and Industrial Institute, Next Generation Eco-Innovation Project (No:413-111-003). Also, the authors acknowledged the editor and anonymous reviewers for their constructive comments.

References

American Public Health Association (APHA) (2005), *Standard Methods for the Examination of Water and Wastewater,* 21st ed., APHA, Washington, DC, U.S.A.

APHA (2009), *Standard Methods for the Examination of Water and Wastewater*, Revised online ed., APHA, Washington, DC, U.S.A.

Brinkmeyer, R. (2016), "Diversity of bacteria in ships ballast water as revealed by next generation DNA sequencing", *Marine Pollut. Bull.,* **107**(1), 277-285.

Bushon, R., Brady, A., Christensen, E. and Stelzer, E. (2018), "Multi-year microbial source tracking study characterizing fecal contamination in an urban watershed", *Water Environ. Res.,* **89**(2), 127-143.

Cabral, A., Stark, J., Kolm, H. and Martins, C. (2018), "An integrated evaluation of some faecal indicator bacteria (FIB) and chemical markers as potential tools for monitoring sewage contamination in subtropical estuaries", *Environ. Pollut.,* **235**, 739-749.

Eren, A.M., Sogin, M. and Maignien, L. (2016), "Editorial: New insights into microbial ecology through subtle nucleotide variation", *Front. Microbiol.,* **7**, 1318.

Galfi, H., Österlund, H., Marsalek, J. and Viklander, M. (2016), "Indicator bacteria and associated water quality constituents in stormwater and snowmelt from four urban catchments", *J. Hydrol.,* **539**, 125-140.

Huang, H., Chow, C. and Jin, B. (2016), "Characterisation of dissolved organic matter in stormwater using high-performance size exclusion chromatography", *J. Environ. Sci.,* **42**, 236-245.

Ibekwe, A., Ma, J. and Murinda, S. (2016), "Bacterial community composition and structure in an urban river impacted by different pollutant sources", *Sci. Total Environ.,* **566-567**, 1176-1185.

Korean Ministry of Environment (2011), *Korean Standard Methods for the Examination of Drinking Water*, Notice 2011-2021, Korea Ministry of Environment, Republic of Korea.

Leung, H., Chen, G. and Sharma, K. (2005), "Effect of detached/re-suspended solids from sewer sediment on the sewage phase bacterial activity", *Water Science and Technology,* **52**(3), 147-152.

Liao, H., Krometis, L.H, Cully Hession, W., Benitez, R., Sawyer, R., Schaberg, E. von Wagoner, E. and Badgley, B.D. (2015), "Storm loads of culturable and molecular fecal indicators in an inland urban stream", *Sci. Total Environ.,* **530-531**(15), 347-

365.
Mcllelan, S., Fisher, J. and Newton, R. (2015), "The microbiome of urban waters", *Int. Microbiol.,* **18**, 141-149.
Ministry of the Environment, Korea (MOE) (2014), *Yearbook of Statistics Korea*, MOE, Republic of Korea.
Paule, M.A., Memon, S.A., Lee, B.Y., Umer, S.R. and Lee, C.H. (2014), "Stormwater runoff quality in correlation to land use and land cover development in Yongin, South Korea", *Water Sci. Technol.*, **70**(2), 218-225.
Paule-Mercado, M.A., Ventura, J.S., Memon, S.A., Jahng, D., Kang, J.H. and Lee, C.H. (2016), "Monitoring and predicting the fecal indicator bacteria concentrations from agricultural, mixed land use and urban stormwater runoff", *Sci. Total Environ.,* **550**, 1171-1181.
Shanks, O.C., Newton, R.J., Keltya, C.A., Huse, S.M., Sogin, M.L. and McLellan, S.L. (2013), "Comparison of the microbial community structures of untreated wastewaters from different geographic locales", *Appl. Environ. Microbiol.,* **79**(9), 2906-2913.
Staley, Z., Grabuski, J., Sverko, E. and Edge, T. (2016), "Comparison of microbial and chemical source tracking markers to identify fecal contamination sources in the humber river (Toronto, Ontario, Canada) and associated storm water outfalls", *Appl. Environ. Microbiol.,* **82**(21), 6357-6366.
Ulrich, N., Rosenberger, A., Brislawn, C., Wright, J., Kessler, C., Toole, D., Solomon, C., Strutt, S., McClure, E. and Lamendella, R. (2016), "Restructuring of the aquatic bacterial community by hydric dynamics associated with superstorm sandy", *Appl. Environ. Microbiol.,* **82**(12), 3525-3536.
Unno, T., Jang, J., Han, D., Kim, J.H., Sadowsky, M.J., Kim, O.S., Chun, J. and Hur, H.G. (2010), "Use of barcoded pyrosequencing and shared OTUS to determine sources of fecal bacteria in watersheds", *Environ. Sci. Technol.,* **44**(20), 7777-7782.
United States Environmental Protection Agency (USEPA) (2012), *Recreational Water Quality*, USEPA, Ohio, U.S.A.

CC