Impact of chemical and microbiological water quality on bacterial community assemblage of San Juan River (Sierra del Rosario, Biosphere Reserve, Cuba)

Impacto de la calidad químicay microbiológica del agua sobre la estructura de la comunidad bacteriana del río San Juan (Sierra del Rosario, Reserva de la Biosfera, Cuba)

Jeny A. Larrea Murrell\textsuperscript{1,2}
Tristano Bacchetti De Gregoris\textsuperscript{2}
Mayra Heydrich Pérez\textsuperscript{1}
Daysi Lugo Moya\textsuperscript{1}
Abraham Estéve Núñez\textsuperscript{2}
Karina Boltes Espínola\textsuperscript{2}
Marcia M. Rojas Badía\textsuperscript{1}

\textsuperscript{1}Laboratorio de Ecología Microbiana, Departamento de Microbiología y Virología, Facultad de Biología, Universidad de La Habana, 25 No. 455, entre J e I Vedado. Habana. Cuba. adina@fbio.uh.cu (https://orcid.org/0000-0002-6067-6331), mayra@fbio.uh.cu, daysi@fbio.uh.cu, marcia@fbio.uh.cu (https://orcid.org/0000-0003-4958-1873)

\textsuperscript{2}Instituto Madrileño de Estudios Avanzados Agua (IMDEA Agua). C/Punto Net 4, Parque Científico Tecnológico de la Universidad de Alcalá.
El objetivo del presente estudio fue analizar la influencia de diferentes indicadores fisicoquímicos y microbiológicos de estructuras de la comunidad bacteriana en un ecosistema de agua dulce ubicado en la Reserva de la Biosfera Sierra del Rosario. La composición de las comunidades de bacterioplancton se determinó por electroforesis en gradiente desnaturizante (DGGE) del 16S rDNA durante tres campañas en tres estaciones de muestreo en el río San Juan. Se realizaron análisis estadísticos multivariados de los patrones de la electroforesis, en relación con parámetros químicos, biológicos y físicos. Los análisis mostraron la variación estacional de la comunidad bacteriana en las estaciones de muestreo menos contaminadas. Además, los sólidos disueltos totales (TDS), la concentración de E. coli y el oxígeno disuelto influyeron significativamente en los perfiles de DGGE en el río San Juan, lo que indica que al menos dos de las tres estaciones de muestreo analizadas tuvieron una buena calidad del agua. Con la combinación de variables ambientales y la DGGE, es posible analizar la influencia de la contaminación química y microbiana en la comunidad bacteriana del río San Juan. En los sitios menos contaminados, las bajas concentraciones de TDS y los indicadores bacterianos de contaminación se relacionaron con la diversidad bacteriana. La presente investigación constituye el primer análisis independiente de cultivo de las comunidades bacterianas de ecosistemas dulceacuícolas en una Reserva de la Biosfera en las islas del Caribe.

**Palabras clave:** Electroforesis en gel de gradiente desnaturizante (DGGE), ecosistema de agua dulce, país tropical, análisis multivariado, Reserva de la Biosfera
The objective of the present study was to analyze the influence of different physicochemical and microbiological indicators on bacterial assemblages in a freshwater ecosystem located in a Biosphere Reserve. The bacterioplankton communities composition was determined by denaturing gradient gel electrophoresis (DGGE) of 16S rDNA during three sampling campaigns in three sampling stations of San Juan River. Multivariate statistical analyses of gel patterns, in relation to chemical, biological and physical parameters were performed. The analyses showed a seasonal variation of bacterial community at the less polluted stations. The correspondence canonical analysis showed that total dissolved solids (TDS), E. coli concentration and dissolved oxygen were significantly influencing the DGGE profiles at San Juan River, indicating that at least two of the three sampling stations analyzed possessed good water quality. With the combination of environmental variables and the DGGE molecular data it could be appreciated the influence of chemical and microbial contamination on the bacterial community of San Juan River. At the less polluted stations low concentrations of TDS and bacterial indicators of fecal contamination were related with bacterial diversity. The present investigation constitutes the first culture-independent analysis of freshwater bacterial communities in a Biosphere Reserve of the Caribbean islands.

**Keywords**: Denaturing gradient gel electrophoresis (DGGE), freshwater ecosystem, tropical country, multivariate analysis, Biosphere Reserve

Received: 17/07/2019
Accepted: 10/08/2020

Introduction
Bacterial communities are highly sensitive to environmental changes as well as to persistent changes in ecosystems due to pollution from human activity or global climate change (Lear et al., 2008; Washington et al., 2013). Recent research on bacterial communities suggests that they may offer an additional tool for assessing the ecological quality of watercourses (Lear et al., 2012; Washington et al., 2013; Newton and McLellan, 2015; Wang et al., 2016). This is due to the advantages offered by bacterial communities as indicators, as they may experience changes in their composition in response to environmental disturbances, the effects of these changes being detected before they are transmitted to higher levels of trophic chains in aquatic ecosystems (Lewis et al., 2010). In addition, they can be sampled in sites where other bioindicators such as benthic invertebrates and fish are not present (Lear et al., 2012).

Culture-independent methods contribute to a reduction of bias in the analysis of the richness of bacterial communities compared to culture-dependent methods, due to the selective pressure imposed by the latter to grow in a solid culture medium, favoring the growth of those populations within the bacterial community capable of forming colonies. It is for this reason that culture-dependent methods underestimate species richness in a sample, which is best estimated through culture-independent methods that can detect viable but non-culturable bacteria (Burtscher et al., 2009; Vaz-Moreira et al., 2011).

Molecular fingerprinting methods such as Denaturing Gradient Gel Electrophoresis (DGGE), the Terminal Restriction Fragment Length Polymorphism (T-RFLP) and the Automated Ribosomal Intergenic Spacer Analysis (ARISA), allow the establishment of community profiles and phylogenetic affiliations of the detected species in case the fragments of DNA can be sequenced. Particularly the DGGE, allows obtaining taxonomic information, because the bands obtained can be split, reamplified and sequenced. Thus, this technique combines the
advantages of cloning, sequencing and T-RFLP and can be used for the analysis of bacterial communities of different ecosystems (Duarte et al., 2012; Liu et al., 2012).

The hydromineral sector "Las Terrazas" is located in the municipality of La Palma, Artemisa province, on the banks of the San Juan River, within the development area of Las Terrazas Resort, in the northeastern part of the Sierra del Rosario Biosphere Reserve. The waters of the sector are representative of the saturation zone of the carbonated massifs and emerge in the contact with parallel and transverse faults to the Falsa Pinar (Peña et al., 2001). Taking into account global mineralization, waters are considered mineral-medicinal and according to their use, they classify as waters of balneological and therapeutic use (Peña et al., 2001). Although the evaluation of the chemical and microbiological quality of the San Juan River (Arpajón et al., 2012; Larrea et al., 2013; Romeu et al., 2015) has been carried out before, culture-independent methods have never been implemented to evaluate the bacterial community of this ecosystem. This type of research could be beneficial to know from the synergistic point of view the influence of different physicochemical variables on the structure of bacterial communities, which can contribute to the improvement of the management of this freshwater ecosystem, which has a great importance for tourism and landscape viewpoint. This study aims to analyze the influence of different physicochemical and microbiological indicators on bacterial assemblages in a freshwater ecosystem located in a Biosphere Reserve.

Materials and methods

Sampling and environmental data
To analyze the bacterioplankton community structure at San Juan River, sampling was done on three times (November 2011; February 2012 and April 2012) at three sampling stations: Presa El Palmar (S-1) (22°50'42.35'' N and 82°56'23.33'' W), Presa San Juan (S-2) (22°50'46.01'' N and 82°56'27.09'' W) and Baños del San Juan (S-3) (22°49'24.02'' N and 82°55'35.08'' W).

For sample collection, sterile plastic sampling bottles (1 L) were used (APHA, 2012). All samples were collected by triplicates. Water samples were taken about one meter from the shore and 15 cm deep. Field parameters consisting in electrical conductivity, pH, temperature, total dissolved solid (TDS) and dissolved oxygen (DO) were measured at each sample location with the use of a multimeter (Mettler Toledo). All water samples were transported on ice to the laboratory and analyzed within 12 h.

Subsamples of water used for nutrients analysis (PO$_4^{3-}$-P, NH$_4^+$-N, NO$_2^-$/N, and NO$_3^-$-N) and chemical oxygen demand (COD) were filtered with oven-dried GF/F filters (Whatman International Ltd., Maidstone, United Kingdom) according to the standard French protocols AFNOR (2009). The determination of inorganic nutrients and COD was carried out using colorimetric methods. Specifically, COD was determined by the dichromate reactor digestion method, for which the COD cuvette test (Spectroquant 1,152,0001, range 0-1500 mg.L$^{-1}$ COD) was used. For the determination of ammonia, the Nessler method was used, and for phosphate, the blue ascorbic acid method was used. All these methods are described in American Public Health Association (APHA)(2012) Precipitations data were obtained from the Ecological station in Artemisa province, Cuba.
Determination of thermotolerant coliforms, *Escherichia coli* and enterococci

Thermotolerant coliforms (TtC) were quantified by membrane filtration (0.45 µm-pore-size, 47-mm-diameter sterile cellulose nitrate filters, Sartorius) on Lactose agar with Tergitol (0.095 ‰ wt/vol final concentration) and triphenyl 2,3,5-tetrazolium chloride (TTC) (0.024 ‰ wt/vol final concentration) (AFNOR, 2009), *Escherichia coli* (EC) and enterococci (Ent) were quantified by the same method on Coliform Chromocult agar and Enterococci Chromocult agar respectively according to Ouattara et al. (2014). The bacteria were cultivated 24 h at 37 °C for EC and Ent and 24 h at 44 °C for TtC. Orange colonies producing a yellow halo under the membrane after incubation were considered thermotolerant coliforms colonies, dark-blue colonies were considered *E. coli* colonies and red colonies were considered enterococci. Counts were expressed as colony-forming units (CFU) per 100 mL of sample.

Analysis of bacterial communities of San Juan River

*DNA extraction and amplification of 16S rDNA fragments*

Microbial biomass was collected from the water column and concentrated by filtration. Aliquots of each sample (1 L) were filtered in triplicate through cellulose nitrate filters of 5 µm and consequently through cellulose nitrate filters of 0.22 µm. From 0.22 µm filters, DNA extraction was performed by SDS-phenol-chloroform method (Oh et al.,
2011). All DNA samples were purified by DNeasy Blood and Tissue Kit (Qiagen).

In order to increase the sensitivity and to facilitate DGGE analysis, nested PCR technique was performed. During the first amplification the PCR tube containing 25 µL of the reaction mixture consisting of a concentration of 1 X of OneTaq Quick-Load 2 X Master Mix mixed with standard buffer (BioLabs, New England), a final concentration of 0.2 µM of the universal primers 27F (5´-AGAGTTTGATCMTGGCTCAG-3´) and 1492R (5´-TACCTTGTTACGACTT-3´) for amplifying 16S rDNA gene (Zhou et al., 2009), 1 µL template DNA (~70 ng) and sterile MilliQ water. After an initial denaturation step of 5 min at 94 °C, samples were amplified for 30 cycles of 30 s at 94 °C, 30 s at 55 °C and 90 s at 72 °C and one final extension step of 5 min at 72 °C.

From the products of the first amplification reaction, two independent amplification reactions were performed, one with primers 341FGC clamp (5´-
GCCCCGCCGCGCGCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGCTACGGGAGG-3´) and 534R (5´-ATTACCGCGGTGCTCTGG-3´) which amplify the V3 variable region of 16S rDNA of Bacteria domain (Muyzer et al., 1993) and the other with the primers 517FGC (5´-GTGCCAGCGCAGCGCCG-3´) (Gich et al., 2005) and 1061R (5´-CAGCAGCAGCTGAGCAG-3´) (Degnan and Ochman, 2012) which amplify the V4-V6 variable regions of 16S rDNA of Bacteria domain. Both reactions were performed in a final volume of 50 µL of the reaction mixture containing a concentration of 1 X Quick-OneTaq of Load Master Mix with 2 X standard buffer (BioLabs, New England), 0.2 µM of the primers, 1 µl of the amplified product and sterile MilliQ water. After an initial denaturation step of 5 min at 95 °C, samples were amplified for 30 cycles of 1 min at 95 °C, 1 min at 53 °C and 2 min at 72 °C and a final extension step of 7 min at 72 °C. All PCR reactions were performed in a Thermal Cycler (ThermalCycler C 1000 TM BIO-RAD).

**DGGE**
DGGE was performed using the system Ingeny phor U (Rafer, Spain). The PCR products obtained from the reaction with primers 341FGC and 534R were loaded onto a 10% polyacrylamide gel (w/v) in TAE 1X buffer (20 mM Tris, 10 mM acetate, 0.5 mM EDTA pH 8), while the products obtained from the reaction with primers 517FGC and 1061R were loaded onto a 7 % polyacrylamide gel. Polyacrylamide gels were made with a denaturing gradient of 30-60 % (where 100 % denaturant containing 7 M urea and 40 % formamide). All electrophoresis were performed at a voltage of 200 V for 4 h and 30 min at 60 °C. Gels were stained with SYBR Gold (1:10 000 dilution) for 30 min. Stained gels were photographed on a UV transilluminator cabinet with a video camera (Alpha Innotech Multimage Light Cabinet Positions Filter).

**Statistical analyses**

Based on the presence (1) or absence (0) of individual bands in each lane, a binary matrix was constructed using the program PyElph 1.4 (Pavel and Vasile, 2012), and a comparative analysis of DGGE profiles was carried out with PRIMER 6 software (PRIMER-E, Ltd., United Kingdom) (Clarke and Gorley, 2006). Cluster analysis using Sorensen similarity coefficient and average linkage technique by the method of unweighted pair group method with arithmetic mean (UPGMA) was performed.

To investigate the relationships between bacterioplankton community and environmental variables in San Juan River, a direct-gradient approach was used, i.e. a canonical correspondence analysis (CCA) using the software package CANOCO, version 4.5 for Windows (ter Braak and Šmilauer, 2002). This method was chosen after initial analysis by detrended correspondence analysis (DCA), which revealed
that the data exhibited a unimodal response to the environmental variables. Data were imported using the WCanoImp program within the CANOCO package. Then was used the CANOCO program to perform CCA with species scaling on intersample distances so that samples and environmental variables formed a biplot. Explanatory variables were added until further addition of variables failed to contribute significantly \((p<0.05)\) to a substantial improvement to the model’s explanatory power. To statistically evaluate the significance of the first canonical axis and of all canonical axes together, a Monte Carlo permutation test was used with 999 unrestricted permutations under the reduced model.

Finally, to represent biplots, the program CANODRAW was used within the CANOCO package. Spearman’s rank pairwise correlations between the different sets of tested environmental variables mentioned above were performed before CCA in order to remove co-varying variables.

**Calculation of Water Quality Index (WQI)**

The Water Quality Index (WQI) was calculated according to Yisa et al. (2010). First the quality rating scale for each parameter \(q_i\) was calculated by using this expression:

\[
q_i = \frac{C_i}{S_i} \times 100
\]

where \((C_i)\) is the concentration of each parameter in each water sample and \((S_i)\) is the standard permissible value for each indicator given by W.H.O. (2003). Then the relative weight \((W_i)\) was calculated by a value inversely proportional to the recommended standard \((S_i)\) of the corresponding parameter:

\[
W_i = \frac{1}{S_i}
\]

Finally, the overall WQI is calculated:
Overall WQI = \( \Sigma q_iW_i / \Sigma W_i \)

According to Ramakrishnaiah et al. (2009) there are five categories for water quality: excellent water quality (<50), good water quality (50-100), poor water quality (100-200), very poor water quality (200-300) and unsuitable (>300).

**Results**

**Physicochemical and microbiological water quality**

Physicochemical and microbiological water quality of San Juan River in the period 2011-2012 is listed in Table 1. Based on their electrical conductivity, phosphate and nitrate concentrations, San Juan River sample stations can be considered mesotrophic according to the classification of Sigee (2005). All sampling stations had good microbiological water quality according to the Cuban guidelines (Norma Cubana 22, 1999). However, Baños del San Juan station (S-3) had high values of nitrate and phosphate concentrations compared with the other two stations. According to the WQI, the water of the stations S-1 and S-2 can be classified as excellent water quality (Table 2). Nevertheless, the water from station S-3 it is classified as unsuitable. Temperature negative correlated with TDS \( (r = -0.71, p<0.05) \) and ammonium \( (r = -0.85, p<0.01) \). pH positive correlated with nitrate \( (r = 0.72, p<0.05) \) and negative correlated with COD \( (r = -0.84, p<0.01) \). Electrical conductivity showed a positive correlation with TDS \( (r = 0.97, p<0.01) \) and negative correlated with precipitations \( (r = -0.77, p<0.05) \).
Dissolved oxygen negative correlated with precipitations \((r = -0.82, p<0.01)\) and nitrite concentration positive correlated with phosphates concentration \((r = 0.81, p<0.01)\).

**Table 1.** Environmental data recorded from San Juan River sampling stations.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature (°C)</td>
<td>25.2</td>
<td>24.9</td>
<td>28.4</td>
<td>24.9</td>
<td>24.6</td>
<td>27.9</td>
<td>25.0</td>
<td>22.6</td>
<td>25.0</td>
</tr>
<tr>
<td>pH</td>
<td>8.39</td>
<td>8.38</td>
<td>8.63</td>
<td>8.60</td>
<td>8.38</td>
<td>8.81</td>
<td>8.90</td>
<td>8.27</td>
<td>8.63</td>
</tr>
<tr>
<td>Conductivity (µs.cm(^{-1}))</td>
<td>450</td>
<td>360</td>
<td>300</td>
<td>450</td>
<td>360</td>
<td>350</td>
<td>480</td>
<td>430</td>
<td>406</td>
</tr>
<tr>
<td>TDS (mg.L(^{-1}))</td>
<td>310</td>
<td>240</td>
<td>200</td>
<td>300</td>
<td>240</td>
<td>230</td>
<td>320</td>
<td>310</td>
<td>301</td>
</tr>
<tr>
<td>DO(mg.L(^{-1}))</td>
<td>11.32</td>
<td>6.97</td>
<td>7.15</td>
<td>9.60</td>
<td>4.37</td>
<td>6.96</td>
<td>9.36</td>
<td>6.82</td>
<td>6.47</td>
</tr>
<tr>
<td>COD (mg.L(^{-1}))</td>
<td>80.0</td>
<td>115.0</td>
<td>90.0</td>
<td>53.3</td>
<td>148.3</td>
<td>60.0</td>
<td>45.0</td>
<td>125.0</td>
<td>61.7</td>
</tr>
<tr>
<td>NO(_3)-N (mg.L(^{-1}))</td>
<td>2.73</td>
<td>5.84</td>
<td>9.45</td>
<td>15.65</td>
<td>6.87</td>
<td>10.16</td>
<td>14.60</td>
<td>2.20</td>
<td>28.07</td>
</tr>
<tr>
<td>NO(_2)-N (mg.L(^{-1}))</td>
<td>0.02</td>
<td>0.02</td>
<td>0.03</td>
<td>0.04</td>
<td>0.03</td>
<td>0.02</td>
<td>0.06</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>NH(_4)-N (mg.L(^{-1}))</td>
<td>0.34</td>
<td>0.31</td>
<td>0.003</td>
<td>0.25</td>
<td>0.55</td>
<td>0.003</td>
<td>0.48</td>
<td>0.61</td>
<td>0.09</td>
</tr>
<tr>
<td>PO(_4)-P (mg.L(^{-1}))</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
<td>0.05</td>
<td>0.02</td>
<td>0.02</td>
<td>1.27</td>
<td>0.13</td>
<td>0.13</td>
</tr>
<tr>
<td>ThC.100 mL(^{-1})</td>
<td>4.0(\times)10(^1)</td>
<td>2.6(\times)10(^1)</td>
<td>9.3(\times)10(^1)</td>
<td>1.3(\times)10(^1)</td>
<td>3.3(\times)10(^2)</td>
<td>5.0(\times)10(^1)</td>
<td>1.1(\times)10(^2)</td>
<td>6.0(\times)10(^1)</td>
<td>5.0(\times)10(^2)</td>
</tr>
<tr>
<td>EC.100 mL(^{-1})</td>
<td>0</td>
<td>4.3(\times)10(^1)</td>
<td>6.7(\times)10(^0)</td>
<td>0</td>
<td>0</td>
<td>1.0(\times)10(^1)</td>
<td>1.1(\times)10(^2)</td>
<td>3.3(\times)10(^1)</td>
<td>6.3(\times)10(^1)</td>
</tr>
</tbody>
</table>
ThC: Thermotolerant coliforms; EC: *Escherichia coli*; Ent: Enterococci; TDS: Total Dissolved Solids; COD: Chemical Oxygen Demand; DO: Dissolved Oxygen; S-1: Presa El Palmar station; S-2: Presa San Juan station; S-3: Baños del San Juan station.

Table 2. Computed WQI values for San Juan River.

<table>
<thead>
<tr>
<th>Environmental data</th>
<th>Standard value (Si)</th>
<th>Wi</th>
<th>qi</th>
<th>qiWi</th>
<th>Overall WQI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>S-1</td>
<td>S-2</td>
<td>S-3</td>
</tr>
<tr>
<td>pH</td>
<td>6.50-8.50</td>
<td>0.13</td>
<td>113.33</td>
<td>114.67</td>
<td>114.67</td>
</tr>
<tr>
<td>Conductivity</td>
<td>20.00-1500.00</td>
<td>0.00</td>
<td>49.30</td>
<td>52.60</td>
<td>58.30</td>
</tr>
<tr>
<td>TDS</td>
<td>500.00</td>
<td>0.00</td>
<td>51.00</td>
<td>53.00</td>
<td>62.10</td>
</tr>
<tr>
<td>DO</td>
<td>4.00</td>
<td>0.25</td>
<td>228.60</td>
<td>174.60</td>
<td>197.90</td>
</tr>
<tr>
<td>COD</td>
<td>10.00</td>
<td>0.10</td>
<td>975.00</td>
<td>1008.00</td>
<td>850.00</td>
</tr>
<tr>
<td>NO₃-N</td>
<td>10.00</td>
<td>0.10</td>
<td>60.90</td>
<td>112.60</td>
<td>151.35</td>
</tr>
<tr>
<td>NO₂-N</td>
<td>0.50</td>
<td>2.00</td>
<td>5.00</td>
<td>6.00</td>
<td>10.00</td>
</tr>
<tr>
<td>NH₄-N</td>
<td>0.50</td>
<td>2.00</td>
<td>34.30</td>
<td>55.30</td>
<td>70.00</td>
</tr>
<tr>
<td>PO₄-P</td>
<td>0.10</td>
<td>10.0</td>
<td>25.00</td>
<td>35.00</td>
<td>700.0</td>
</tr>
</tbody>
</table>

The data were obtained from the Ecological Station Sierra del Rosario, Artemisa. It was taking into account the precipitations of 10 days before the sampling date.
**Analysis of bacterial communities from San Juan River by DGGE**

The Fig. 1a shows the pattern of bands obtained by the DGGE from the San Juan River samples. The number of bands per sample varied between 15 (S-3 November 2011) and 26 (S-2 April 2012), in addition it is possible to observe differences in the bands profiles of each sampling station in each of the analyzed months; thus as differences in the intensity of the bands.

Taking into account the bands profiles obtained by DGGE, a clustering analysis was performed (Fig. 1b), where the formation of three groups with 60 % similarity was observed. The first group consisting of samples of S-3 station collected in February and April 2012, the second group includes samples from the S-1 sampling station collected in the months of February and April 2012, and the third one groups the samples of the S-2 station collected in February and April 2012. The samples of the month of November of each one of the stations analyzed were separated from each other. The permutation analysis of variance (PERMANOVA) indicated significant differences between the bacterial communities of the different sampling stations ($p = 0.001$), as well as between the different sampled months ($p = 0.001$). This result can be seen in the cluster analysis where each group formed is represented by a particular station, in addition there is a separation between the samples collected
in the months February and April 2012 and the samples obtained in November 2011.

**Figure 1.** DGGE analysis of 16S rDNA fragments amplified by PCR with primers 341FGC and 534R. (A) Image of the gel obtained from samples of San Juan River collected in the period November 2011-April 2012. Denaturant gradient used: 30-60%. (B) Cluster analysis by UPGMA method based on band patterns of samples from San Juan River. The similarity was expressed as percentage of Sorensen index. San Juan River sampling stations: S-3: Baños del San Juan; S-2: Presa San Juan; S-1: Presa El Palmar.

**Analysis of the relationship between the bacterial community assemblage and the physicochemical and fecal contamination indicators**
Total dissolved solids ($p=0.015$) and *E. coli* concentration ($p=0.045$) were the main factors that influenced the variability of the bacterial community assemblage in San Juan River (Figure 2, Table 3). Monte Carlo test for all canonical axes ($p=0.002$) was significant, indicating that the selected variables can explain the variability of the bacterial community structure in San Juan River stations. Total dissolved solids, *E. coli* concentration and dissolved oxygen, explained 21.2 %, 19.4 %, and 17.3 % respectively, of the bacterioplankton variability in San Juan River. According to the canonical correspondence analysis, the cumulative percentage of variance explained by the species-environmental variables indicated that the first and second canonical axes describe 76.8 % of the variance. Correlations species-environment from each axis, were 0.93 and 0.98 for the first and second axis respectively. Dissolved oxygen and TDS were the two main components of the first axis, and *E. coli* concentration was the main component of the second axis. Based on the pattern in the biplot, we divided the samples into three clusters: Cluster 1, most of the samples from stations S-1 and S-2 (not including November 2011) were characterized by the lowest values of TDS and *E. coli* concentration; Cluster 2, samples from station S-3 (from November 2011 to April 2012) were characterized by the highest values of TDS and *E. coli* concentration and Cluster 3, included samples from stations S-1 and S-2 (November 2011) characterized by the highest values of dissolved oxygen.
Figure 2. Biplot of canonical correspondence analysis (CCA) of the bacterial community assemblage from San Juan River in the period November 2011-April 2012 using physicochemical and microbiological indicators. The red arrows indicate the direction in which the value of each indicator of water quality increases. The length of the arrows indicates the degree of correlation with the axis shown. The position of the sampling stations on the arrows is interpreted by the projection of the symbols on the arrows, and indicates the degree of influence of the water quality indicator represented by the red arrow on the bacterial community assemblage of the sampling station. DO: Dissolved Oxygen; TDS: Total Dissolved Solids; EC: *Escherichia coli*; S-1: Presa El Palmar station; S-2: Presa San Juan station; S-3: Baños del San Juan station.

Table 3. Summary of results from canonical correspondence analyses (CCA) of the bacterioplankton community structure data when constrained by water quality indicators in San Juan River.

|        | Axis 1 | Axis 2 | Axis 3 | Total Variance |
|--------|--------|--------|--------|----------------|----------------|

17
<table>
<thead>
<tr>
<th></th>
<th>0.44</th>
<th>0.35</th>
<th>0.24</th>
<th>1.92</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Eigenvalues</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Species-environment</strong></td>
<td>0.93</td>
<td>0.98</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td><strong>correlations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cumulative percentage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>variance:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>of species data</td>
<td>22.70</td>
<td>41.00</td>
<td>53.30</td>
<td></td>
</tr>
<tr>
<td>of species-environment</td>
<td>42.60</td>
<td>76.80</td>
<td>100.00</td>
<td></td>
</tr>
<tr>
<td>relation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sum of all eigenvalues</strong></td>
<td></td>
<td></td>
<td></td>
<td>1.92</td>
</tr>
<tr>
<td><strong>Sum of all canonical</strong></td>
<td></td>
<td></td>
<td></td>
<td>1.02</td>
</tr>
<tr>
<td><strong>eigenvalues</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Correlations with the</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>axes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dissolved oxygen (DO)</td>
<td>0.56</td>
<td>0.53</td>
<td>-0.63</td>
<td></td>
</tr>
<tr>
<td>Total dissolved solids</td>
<td><strong>0.89</strong></td>
<td>0.32</td>
<td>0.31</td>
<td></td>
</tr>
<tr>
<td>(TDS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Escherichia coli</strong></td>
<td>0.60</td>
<td>-0.75</td>
<td>-0.29</td>
<td></td>
</tr>
<tr>
<td>concentration</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Bold type indicates the strong factors correlated with the first two CCA ordination axes 
($r > 0.70$ or $r < -0.70$).

**Discussion**

**Physicochemical and microbiological quality of San Juan River**

18
In the context of assessing the impact of pollutants on the diversity of microbial communities in freshwater ecosystems, the difference between the relative effects of selective pressure due to contamination, and those that result from natural environmental factors and processes is very important. Natural variations can be found in the structure and function of non-impacted communities when different ecosystems, or different sampling stations within the same ecosystem, are compared as refer Dorigo et al. (2009). That is for; polluted and non-polluted stations in San Juan River are evaluated to analyze the influence of environmental parameters on bacterial community structure in the present study.

San Juan River is located in a rural area and particularly in the Biosphere Reserve Sierra del Rosario where an active Sustainable Development Program contributes to the environment protection (Larrea et al., 2013). Waters in this environment are very clean with abundant fishes, some macrophytes and good odor. Although most of the physicochemical and microbial indicators do not exceed the Cuban guidelines for bathing sites in coastal and inland water bodies (Norma Cubana 22, 1999); S-3 station presented the highest values of the microbial and chemical variables, compared with the other stations. This fact could be possible because of the anthropogenic contribution generated by a main affluence of public to these natural swimming pools and the effect from the runoff from an oxidation lagoon, which is reflected in the ammonium and fecal indicators concentrations. Considering the WQI in the period evaluated the sampling stations S-1 and S-2 presented excellent water quality compared to the S-3 station, which is in correspondence with studies previously done by Larrea et al. (2013). In order to simplify, inform and interpret the complex information obtained from a water body, different authors have calculated water quality indexes (Yisa and Jimoh, 2010; Rubio-Arias et al., 2012; Rubio-Arias et al., 2016). Thus, with a single number it can be interpreted the level of pollution that a water body presents.
Based on the correlation coefficients obtained in the present study, the temperature was negatively related to total dissolved solids and ammonium concentrations, which could be because the increase in temperature can raise the metabolic rate of microorganisms, leading to degradation of organic matter, which is used as a source of carbon and energy (Dinsmore et al., 2013). This causes the reduction of organic matter in the water column and therefore of total dissolved solids. On the other hand, temperature increase can favor the activity of the nitrifying bacteria, which can use the ammonium in the nitrification process, which consists of two steps, (1) the oxidation of ammonium to nitrite by archaea and ammonium oxidizing bacteria and (2) nitrite oxidation to nitrate carried out by nitrite oxidizing bacteria (Chen et al., 2010). In this way, with temperature increase, would decrease the ammonium concentrations in the waters in the conditions of the San Juan River in the analyzed period.

In the case of electrical conductivity and TDS, both are an indirect measure of the salinity of a solution (Niekerk et al., 2014), which explains the positive correlation observed between both indicators in the river. Electrical conductivity is often used as a substitute for total dissolved solids and is a measure of the ability of a solution to conduct electric current between two points (Dunlop et al., 2005; Niekerk et al., 2014). In contrast, TDS provide a general indication of the level of dissolved solids less than two micrometers present in a fluvial stream, which includes all dissociated electrolytes that increase the concentration of salts and other compounds such as dissolved organic matter (Tanor et al., 2014). In addition, a negative correlation between electrical conductivity and precipitation was observed, which may be due to the fact that the electrical conductivity tends to change as a function of the ecosystem recharge: during rainy periods it tends to decrease its value due to the salts dilution; while during the dry periods the opposite occurs (Rashed and Younis, 2012).

COD was negatively related to the pH, which may be because the greater the amount of organic matter in the ecosystem, the greater
amount of dissolved oxygen needed to decompose it, a process that contributes to the release of carbon dioxide, which decreases the pH of the water. Varol et al. (2012) argue that the decomposition of organic matter leads to acidification and therefore to a decrease in pH.

In the San Juan River, a negative correlation was observed between dissolved oxygen concentrations and precipitation. Similar results were obtained by Jung et al. (2014) on the Han river in South Korea, who observed that precipitation was negatively related to oxygen and positively to the concentration of suspended solids. Rainfall can contribute to increased turbidity in aquatic systems because they cause outcrops of sediment from the river bottom and drag the adjacent soils; this decreases the penetration of sunlight, which affects the phytoplanktonic activity and therefore the process of photosynthesis with the consequent decrease of dissolved oxygen. Therefore an increase in the turbidity of the water may cause a decrease in the process of photosynthesis and therefore a decrease in pH and dissolved oxygen concentrations.

On the other hand, there was a positive correlation between nitrite and phosphate concentrations, which were higher in S-3 station compared to S-1 and S-2 stations. A cause that could contribute to these concentrations of nutrients, are the runoff from the aerobic oxidation lagoon that is located between stations S-2 and S-3, which presents ruptures and receives volumes of domestic wastewater that are superior to its retention capacity, as observed during the sampling. In this same period a positive correlation was observed between the high concentrations of nitrates and pH, results similar to those obtained by Zhang et al. (2014) at Lake Erhai in China. These authors suggest that at pH 8 the growth and metabolism of nitrifying bacteria is favored. Furthermore, under aerobic conditions, the release of ammonia from the sediments decreases and this is converted into nitrate by nitrification. This phenomenon contributes to the reduction of ammonium and the increase of nitrates in the waters.
Relationship between the bacterial community assemblage and the physicochemical and fecal contamination indicators in San Juan River

The combination of DGGE with multivariate methods allowed to appreciate the spatio-temporal variability in the assemblage of bacterial communities during the study period, which corresponds to results obtained by different authors in different freshwater ecosystems using this molecular technique (Berdjeb et al., 2011a; Figueiredo et al., 2012; Ibekwe et al., 2012; Tang et al., 2014).

Based on the structure of bacterial communities, it was observed that the month of November 2011 is separated from the months of February and April 2012 (Fig. 1b). In addition, the less contaminated stations (S-1 and S-2) are separated from the station with the highest contamination (S-3), which supports the idea that bacterial communities can be used as indicators of ecological status of aquatic ecosystems (Lear et al., 2012; Washington et al., 2013). It is argued that bacterial communities are highly sensitive to environmental changes as well as to persistent changes in ecosystems due to pollution from human activity or global climate change (Lear et al., 2008). The temporal and spatial changes detected may be related to abiotic factors such as: temperature, dissolved oxygen, pH, electrical conductivity, water transparency, organic matter concentration, nutrients, among others, and biotics such as: chlorophyll a content, interactions with phytoplankton and zooplankton, predation, parasitism and competition, which regulate community changes in aquatic ecosystems (Lindström et al., 2005; Berdjeb et al., 2011a; Berdjeb et al., 2011b; Llirós et al., 2014).
The factors that most influenced the assemblage of the bacterial communities of the San Juan River in the period November 2011-April 2012 were the TDS and the concentrations of *E. coli*, the latter used as an indicator of fecal contamination. Both variables presented high values at station S-3 compared to stations S-1 and S-2, which influences changes in bacterioplankton. The presence of high concentrations of TDS, denote the contribution of soils through the natural processes of trawling and drainage; but also, the resuspension of the sediment, which occur more frequently during rainy periods (Crim *et al.*, 2012), as happened during the months of February and April 2012, where 86.0 mm and 124.9 mm of precipitation were reported respectively. The input of the soil can contribute not only to the entry of bacteria into the river waters, but also to the entry of organic and inorganic compounds, all of which promotes changes in the structure of bacterial communities, as was observed in the Brazos river (Becker *et al.*, 2016). Moreover, resuspension of the sediment in the water column, also contributes to the changes in the structure of bacterial communities, passing sediment microorganisms and nutrients in the water column. High concentrations of TDS can contribute to the decrease of light in the water column and therefore, to the decrease of photosynthesis. This causes a decrease in the production of dissolved oxygen, which contributes to changes in the structure of the bacterial community (Breton-Deval *et al.*, 2019). On the other hand, high concentrations of TDS can cause an increase in salinity, which brings with it changes in biotic communities, a decrease in diversity and the exclusion of species less tolerant to these conditions (Weber-Scannell and Duffy, 2007). In rural landscapes, such as the Sierra del Rosario where the San Juan River is located, the sources of fecal contamination include: cattle, agricultural runoff and leakage from oxidation ponds, the latter used in wastewater treatment (Crim *et al.*, 2012; Sun *et al.*, 2016). Goto and Yan (2011) argue that the soil is an important reservoir of fecal contamination indicator bacteria, and that the exchange between soil and water during abundant precipitation constitutes a means of transporting fecal bacteria. These bacteria can reach the soil from the depositions of wildlife (eg: birds, deer, wild pigs,
etc.) as well as domestic animals. The sources of contamination mentioned above have an impact on the quality of the waters of the S-3 station and therefore on the changes that occur in the bacterial communities of that sampling station.

This research is of great importance because, for the first time in Cuba, an analysis of the structure of the bacterial community of the waters of a Biosphere Reserve, such as the Sierra del Rosario, is carried out, and is related to the quality of the water, which contributes to the conservation of this resource and its environment. As part of the environmental efforts required for the preservation of water as a natural resource, it is necessary to have quantitative data related to water pollution (CITMA, 2009), so the present study constitutes a contribution for the improvement of the management of aquatic ecosystems. The use of bacterial communities as indicators of water quality is very important because it allows the detection of any disturbance in the ecosystem at an early stage before it is transmitted to higher levels of the trophic chain in aquatic ecosystems (Lewis et al. 2010).

**Conclusions**

With the combination of environmental variables and the DGGE molecular data it could be appreciated the influence of chemical and microbial contamination on the bacterial community of San Juan River. At the less polluted stations low concentrations of TDS and bacterial indicators of fecal contamination were related with bacterial diversity. These results contribute to the diagnosis of the aquatic ecosystems analyzed and constitute the first step for management’s plans on the river. The analysis herein has provided useful information on the role of certain environmental variables that shape community structure in
rivers. Further studies including a deep identification of the taxa that compose bacterial community in the freshwater ecosystem analyzed should be interest.

**Acknowledgments**

This work was supported by the CITMA-PNCB project. The authors want to thanks Dr. Maickel Armenteros and Dr. Jose L. Ponce de León for their assistance with the statistics. Also, the authors want to thanks Fidel Hernández and Jorge Luis Zamora, for the help with the access to the Biosphere Reserve Sierra del Rosario, Artemisa, Cuba.

**References**


