Comparison of bacterial communities in the Solimões and Negro River tributaries of the Amazon River based on small subunit rRNA gene sequences

J.C.C. Peixoto1,2, L. Leomil1, J.V. Souza1, F.B.S. Peixoto2 and S. Astolfi-Filho1

1Laboratório de Tecnologias do DNA, Centro de Apoio Multidisciplinar, Universidade Federal do Amazonas, Manaus, AM, Brasil
2Laboratório de Biologia Molecular e Biotecnologia, Departamento de Ciências Biológicas e da Saúde, Coordenação de Ciências Biológicas, Universidade Federal do Amapá, Macapá, AP, Brasil

Corresponding author: S. Astolfi-Filho
E-mail: spartaco.biote@gmail.com

Received May 24, 2011
Accepted August 19, 2011
Published December 8, 2011
DOI http://dx.doi.org/10.4238/2011.December.8.8

ABSTRACT. The microbiota of the Amazon River basin has been little studied. We compared the structure of bacterial communities of the Solimões and Negro Rivers, the main Amazon River tributaries, based on analysis of 16S rRNA gene sequences. Water was sampled with a 3-L Van Dorn collection bottle; samples were collected at nine different points/depths totaling 27 L of water from each river. Total DNA was extracted from biomass retained by a 0.22-μm filter after sequential filtration of the water through 0.8- and 0.22-μm filters. The 16S rRNA gene was amplified by PCR, cloned and sequenced, and the sequences were analyzed with the PHYLIP and DOTUR programs to obtain the operational taxonomic units (OTUs) and to calculate the diversity and
richness indices using the SPADE program. Taxonomic affiliation was determined using the naive Bayesian rRNA Classifier of the RDP II (Ribosomal Database Project). We recovered 158 sequences from the Solimões River grouped into 103 OTUs, and 197 sequences from the Negro River library grouped into 90 OTUs by the DOTUR program. The Solimões River was found to have a greater diversity of bacterial genera, and greater estimated richness of 446 OTUs, compared with 242 OTUs from the Negro River, as calculated by ACE estimator. The Negro River has less bacterial diversity, but more 16S rRNA gene sequences belonging to the bacterial genus *Polynucleobacter* were detected; 56 sequences from this genus were found (about 30% of the total sequences). We suggest that a more in-depth investigation be made to elucidate the role played by these bacteria in the river environment. These differences in bacterial diversity between Solimões and Negro Rivers could be explained by differences in organic matter content and pH of the rivers.

**Key words:** 16S rRNA gene; Bacterial diversity; Amazon basin; Solimões River; Negro River

**INTRODUCTION**

The Amazon River originates in the Peruvian Andes, is approximately 6868 km in length, and is formed by a network of more than a thousand affluents. The volume of water of the Amazon River represents the combined result of the contribution of the rivers that empty their waters into the channel of the Amazon, giving rise to the largest drainage basin in the world, with an area of 7 million km$^2$ (Sioli, 1976). After coursing 1500 km, the Negro River meets the Solimões River, just below the city of Manaus, and after this confluence, the Amazon River begins. The Negro River is the third major affluent channeling to the Amazon River. However, the Solimões and Negro Rivers are distinct in their geomorphology, pH level, conductivity and amount of suspended matter (Sioli, 1976). White water rivers such as the Solimões are rich in nutrients and suspended matter, while black waters such as the Negro are scarce in both, which can determine the spatial distribution of bacterial communities present in the two environments.

The frozen continents (Brinkmeyer et al., 2003; Crump et al., 2003), the geothermal sources of Yellowstone (Hugenholtz et al., 1998), different oceans (Mullins et al., 1995) and freshwater environments such as lakes (Eiler and Bertilsson, 2004; Warnecke et al., 2004) and rivers (Freese et al., 2006) have been exhaustively investigated by a metagenomic approach, aimed at a better understanding of the prokaryotic communities existing in these environments, but still unknown is the role that they play in the respective ecosystems. The isolated study of Borneman and Triplet (1997) on the amplification and sequencing of the SSU rRNA gene described differences in microbial populations between soil from a primary forest and pasture soil in the state of Para, in the Brazilian Amazon region. More recently, Fierer and Jackson (2006) carried out a broad biogeographic analysis of microbial communities in different soils on the American continents, including soil from a forest in the Peruvian Amazon region, relating especially the pH of the soils to the richness of existing species.
The enormity of the Amazon basin accounts for the fact that almost all the microbial diversity present there remains obscure, waiting to be discovered. Fonseca et al. (1982) related total bacterial counts to the autodepuration capacity of the Negro River. Rai and Hill (1978) analyzed the occurrence of bacteria in the Amazon River, close to Manaus, utilizing the standard plate count (SPC) method. However, no studies have been conducted using a metagenomic approach, with the aim of determining more realistically the bacterial diversity of the rivers that make up the Amazon River basin. In this work, we collected water samples and extracted total DNA from the Negro and Solimões Rivers, in the Amazon basin, at locations that precede the confluence of these rivers, to form the Amazon River, with the aim of surveying the bacterial populations present, by the amplification, cloning, sequencing and analysis of the SSU rRNA gene.

MATERIAL AND METHODS

Environmental sampling

The water samples were collected in the months of April and May of 2006 at different points situated in the Solimões and Negro Rivers, namely three points in the Negro and three in the Solimões. The sampling points were at 3 different depths, totaling 9 samples in each river in the channel, and two others about 60 m from the banks, left and right. A 3-L Van Dorn collection bottle was used, where 3 L water were collected at each sampling point, totaling 27 L water from each river. In the laboratory, the water samples were filtered, utilizing a Sterifil® filtration apparatus (Millipore Corp., Bedford, Mass.) connected to a vacuum pump (Quimis). A pre-filtration was performed using glass microfiber filters with a pore size of 0.8 µm, 47 mm in diameter (Durapore; Millipore Corp., Bedford, MA). The water was again filtered utilizing a 0.22-µm filters (Millipore Corp., Bedford, MA), which were changed when they were completely blocked. After filtration, the membranes were frozen at -20°C, and total DNA was then extracted.

DNA extraction

For DNA extraction, the filters were cut with a new scalpel into small pieces (approximately 1 cm by 1 cm) and placed in a 50-mL FALCON tube, to which 6 mL of TEN buffer (100 mM Tris-HCl, pH 8.0, 100 mM EDTA, pH 8.0, 150 mM NaCl, 3% CTAB) were added and the tubes vortexed for 2 min. Next, 10% sodium dodecyl sulfate (SDS), lysozyme (20 mg/mL) and proteinase K (20 mg/mL) were added. The contents were then transferred to 1.5-mL microtubes, and DNA was purified by successive steps of the phenol/chloroform method. After precipitation with isopropanol at room temperature (RT), total DNA was washed with 200 µL chilled 70% ethanol. DNA was kept at RT to dry. Next, DNA was resuspended in 100 µL TE (Tris-HCl 10 mM, EDTA 0.1 mM) and stored overnight at 4°C. The quality and quantity of DNA was determined on 0.8% agarose gel electrophoresis, and the DNA was stored at -20°C. The protocol for DNA extraction was according to Fuhrman et al. (1993) with modifications.

Amplification, SSUrDNA library construction and sequencing

The PCR protocol was according to Crump et al. (1999), with modifications, and con-
sisted of a final volume of 25 μL (2.5 mM MgCl₂, 2.5 mM dNTPs; 5 nM of each primer; 2.5 U Taq DNA polymerase; 1X PCR buffer). The primers used were for bacteria 530F (5'-TTG GAGAGTTTGATCCTGGCTC-3') and 1194R (5'-ACGTCATCCCCACCTTCCTC-3'), which bind respectively to conserved regions close to 5 and 1194 bp of the 16S rRNA gene (Simmon et al., 2006). Amplification was carried out in a thermocycler (Thermo). The amplicons were purified using the GFX™ PCR - DNA and Gel Band Purification kit (Amersham Bioscience) before ligation into the cloning vector. Next, a ligation was made between TOPO-TA Cloning Vector® 4.1 (Invitrogen) and the PCR products of the 16S rRNA gene. The amplification reaction for sequencing of the inserts consisted of the preparation of the Pre-Mix (DYEastic ET Dye Terminator Cycle Sequencing Kit for MegaBACE; Amersham Bioscience). The primer used in the reaction was M13 reverse 5'-TCACACAGGAAACAGCTATGAC-3' (0.5 μM). Sequencing was done in a MegaBace capillary sequencer (Amersham Bioscience), according to manufacturer recommendations.

Data analysis

The sequences determined were analyzed by the phred/phrap program (Ewing and Green, 1998) for removal of bases of low quality, keeping only sequences with a quality >20, and then by VecScreen program to remove contaminating sequences belonging to the cloning vector. The analyses were carried out at the bioinformatics site of the Multidisciplinary Support Center (CAM) of the Federal University of Amazonas (http://portal.ufam.edu.br/).

The sequences were aligned by the CLUSTAL W program, version 1.7 (Thompson et al., 1994). The archive generated by alignment was given a compatible file extension with the DNADIST program - contained in the phylogenetic analysis package PHYLIP, version 3.67 (Felsenstein, 2007). A distance matrix was constructed with DNADIST, utilizing the Jukes-Cantor algorithm. The evolutionary matrix was utilized in the analysis with the program DOTUR (Distance Based OTU and Richness Determination) (Schloss and Handelsman, 2005), for the determination of the number of operational taxonomic units (OTUs). Thus, DOTUR grouped sequences with a 1% difference or less (D = 0.01) in the same OTU.

Taxonomic affiliation was carried out by submitting the sequences to the program Bayesian rRNA Classifier in the database of the Ribosomal Database Project (RDP II) (Wang et al., 2007). Classifier contains in its database sequences of 16S rRNA of “type-strains” and sequences with variable regions of little known bacteria. The confidence interval is at least 80%. Sequences with less than 97% similarity in relation to the sequences deposited in RDP II/Genebank were classified as unidentified.

The estimates of richness of species, the OTUs, were calculated by the SPADE program (Species Prediction and Diversity Estimation) (Chao et al., 2005) using the non-parametric methods ACE and CHAO1. Also calculated were the Shannon diversity index and Simpson’s reciprocal index, besides the estimated sample coverage.

RESULTS AND DISCUSSION

A total of 158 sequences were obtained from the library of the Solimões River and 197 sequences from the library of the Negro River; they were considered of good quality, with a mean size of 495 bp. DOTUR grouped the 158 sequences from the Solimões River into 103
OTUs, and the 197 sequences from the Negro River into 90 OTUs. Each OTU was submitted to the Classifier program of RDP II, to determine the taxonomic affiliation. Of the 103 OTUs from the Solimões library, 93 are related to the phylum Proteobacteria and 6 to the phylum Actinobacteria, and 4 OTUs did not show an affiliation with any known phylum. In the Negro library, 84 OTUs were classified as belonging to the phylum Proteobacteria and 6 to the phylum Actinobacteria (Table 1). The results demonstrate that the library of the Solimões River, with 103 OTUs, showed a greater estimated richness, 337 OTUs, by the Chao1 estimator and 446 OTUs by the ACE estimator, while the Negro River showed a richness estimate of 210 OTUs by the Chao1 estimator and 242 OTUs by the ACE estimator. The Shannon diversity index indicated a superior bacterial diversity in the Solimões River, 4.34 versus 3.97 for the Negro River. The estimated sample coverage (ESC) was 69.5% for the Negro and 48.7% for the Solimões, demonstrating that in order to attain probable total coverage of the OTUs in this environment, it would be necessary to have a sampling of at least double the number of actual sequences (Table 2). The analysis expressed in rarefaction curves indicated greater saturation of OTUs of the Negro River library, compared to the OTUs of the Solimões River (Figure 1).

Many ecological factors considered important, such as pH, water temperature, concentration of organic matter and intensity of ultraviolet radiation (Hader et al., 1998), are responsible for the establishment and propagation of bacterial populations in freshwater microbial communities. The Negro River meets the Solimões River and gives rise to the Amazon River, but the two show contrasting abiotic factors, such as black water in the Negro River and white in the Solimões River, acid pH in the Negro and neutral in the Solimões, and an abundance of organic matter in the Solimões and low concentration in the Negro. Despite pH being noted as a limiting factor in the composition of bacterial communities of soils (Fierer and Jackson, 2006), this does not appear true for aquatic environments. Zwart et al. (2002) showed in a broad analysis of previously published works, that the majority of the sequences of the bacterial SSU rRNA gene, recovered from freshwater environments (rivers and lakes), of different geographic regions and distinct abiotic factors, had a closer phylogenetic relation between them when compared to sequences of soils or marine environments, belonging to the same taxonomic groups. Their analysis also showed that there is a global occurrence of particular bacterial groups in freshwater environments, such as native species, adapting to the abiotic conditions of each location, prompting the authors to devise the term “typical freshwater bacteria.” The same pattern of occurrence was observed in the bacterial communities of the Negro and Solimões Rivers, where both showed sequences of the SSU rRNA gene of bacteria belonging to the classes α-Proteobacteria, β-Proteobacteria, γ-Proteobacteria (present only in the Solimões), and Actinobacteria.

The 16S rRNA gene library of the Negro River showed 56 sequences (16 OTUs), affiliated with the genus Polynucleobacter, which predominated in relation to other genera and even in comparison to the library of the Solimões River, where only 8 sequences (4 OTUs) were classified as Polynucleobacter. The 56 sequences represented about 30% of the total sequences investigated in the Negro River library, which is in agreement with the findings of Hahn et al. (2005), who found in a freshwater lake, with an acid pH, brown coloration, and a high concentration of humic substances, that the genus Polynucleobacter contributed to about 60% of the total prokaryotic cells composing the bacterioplankton. This genus appears to demonstrate an ecological plasticity, referred to by Hahn (2006), which reflects the capacity of particular bacterial taxa to occupy ecological niches with distinct abiotic characteristics, which brings to mind the concept of ecological species (ecotype) proposed by Ward et al. (1998).
<table>
<thead>
<tr>
<th>Environment</th>
<th>Class</th>
<th>Order</th>
<th>Family</th>
<th>Genus</th>
<th>Representative clones</th>
<th>No. of OTUs</th>
<th>No. of sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negro River</td>
<td>Alpha proteobacteria</td>
<td>Rhodospirillales</td>
<td>Aceobacteraceae</td>
<td>Unidentified</td>
<td>NA10-7; NB04-6, NB05-4; NB03-4; NE03-4; NF01-1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Beijerinckiaceae</td>
<td>Unidentified</td>
<td>NA02-1; NA12-6; NB06-6; NB01-1; NE09-7; NF06-7; NF08-7; NF12-1; NG10-2; NG07-6; NH05-6; NH07-6; NH09-1; NC01-4</td>
<td>16</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unidentified</td>
<td></td>
<td>NA02-7; NB03-1; NC03-2; NC06-5; ND01-3; NE11-5; NB12-8; NF06-3; NE11-3; NG12-5; NH04-5; NH06-2; NH10-2; NF08-6</td>
<td>14</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>Betaproteobacteria</td>
<td>Burkholderiales</td>
<td>Burkholderiaceae</td>
<td>Polynucleobacter</td>
<td>NA07-7; NG05-1; NH03-6</td>
<td>3</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unidentified</td>
<td></td>
<td>NA01-5; NA07-2; NA10-3; NA10-6; ND03-5; ND08-7; NE05-7; NE12-2; NE12-4; NF03-3; NF08-2; NG10-5; NA03-4</td>
<td>13</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Incertae sedis 5</td>
<td></td>
<td>NE01-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unidentified</td>
<td></td>
<td>NA12-4; NB05-7; NB08-1; NE07-2; NE07-7; NF05-6; NG01-6; NG10-3; NH10-3; NH02-1; NF11-1; NH07-2; NE05-3; NA06-4; NB10-7</td>
<td>15</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Neisseriales</td>
<td>Rhodocyclales</td>
<td>Rhodocyclaceae</td>
<td>Aquitalea</td>
<td>NA04-5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Denitratisoma</td>
<td>NB01-6</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Unidentified</td>
<td>NC06-2; NF09-1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Continued on next page
<table>
<thead>
<tr>
<th>Environment</th>
<th>Class</th>
<th>Order</th>
<th>Family</th>
<th>Genus</th>
<th>Representative clones</th>
<th>No. of OTUs</th>
<th>No. of sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unidentified</td>
<td>Unidentified</td>
<td>NB02-2; NC10-7; NF07-7; NF08-5; NG04-7; ND06-2; ND04-2; NB02-6; NC03-7; ND06-2; NE04-1</td>
<td>11</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ND12-6; NF07-7; NF08-5; NG04-7; ND06-2; NB02-6</td>
<td>Actinobacteria</td>
<td>Actinomycetales</td>
<td>6</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Subtotal

| | | | | | | 90 | 197 |

Total

| | | | | | | 193 | 355 |

The Classifier program of RDP II was used, from class down to genus. Only *Chromobacterium violaceum* was classified at the species level. Similarity index between 97 and 100%.
In the library of the SSU rRNA gene of the Solimões River, one OTU (SG02-1) containing two sequences was classified with 100% similarity by BLASTn and Classifier (RDPII), in the species *Chromobacterium violaceum*. This bacterium is found in various environments, including rivers and soils, tropical and subtropical (Lima-Bittencourt et al., 2007). In the Negro River, it is widely isolated by plating in culture medium. Curiously, no OTU of this genus was recovered in the library of the SSU rRNA gene of the Negro River, confirming the absence of isolates of *C. violaceum* in the same water samples studied by plate culture in LB medium. The disappearance of isolates of *C. violaceum* in the sites sampled cannot be explained by local abiotic factors, since water samples of the Negro River from different sites showed the growth of violet-colored colonies on culture plates, suggestive of *C. violaceum* (unpublished).

Table 2. Richness of OTUs, Shannon diversity index, Simpson reciprocal (1/D) and ESC.

<table>
<thead>
<tr>
<th>Community</th>
<th>No. seq</th>
<th>No. OTUs</th>
<th>Estimate of richness of OTUs</th>
<th>Diversity index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Chao1</td>
<td>ACE</td>
</tr>
<tr>
<td>Solimões</td>
<td>158</td>
<td>103</td>
<td>337.3 (222.6; 562.2)</td>
<td>446.7 (291.6; 729.3)</td>
</tr>
<tr>
<td>Negro</td>
<td>197</td>
<td>90</td>
<td>210.0 (148.8; 334.9)</td>
<td>242.0 (174.0; 365.2)</td>
</tr>
</tbody>
</table>

No. seq = number of sequences recovered from each environment; No. OTUs = total OTUs determined by the DOTUR program; 1/D = maximal Simpson similarity, reciprocal of the index; ESC = estimated sample coverage.

Figure 1. Rarefaction curves of the operational taxonomic units (OTUs) derived from the libraries of the Solimões (gray line) and Negro Rivers (black line). D = 0.01 (99% similarity).

In the library of the SSU rRNA gene of the Solimões River, one OTU (SG02-1) containing two sequences was classified with 100% similarity by BLASTn and Classifier (RDPII), in the species *Chromobacterium violaceum*. This bacterium is found in various environments, including rivers and soils, tropical and subtropical (Lima-Bittencourt et al., 2007). In the Negro River, it is widely isolated by plating in culture medium. Curiously, no OTU of this genus was recovered in the library of the SSU rRNA gene of the Negro River, confirming the absence of isolates of *C. violaceum* in the same water samples studied by plate culture in LB medium. The disappearance of isolates of *C. violaceum* in the sites sampled cannot be explained by local abiotic factors, since water samples of the Negro River from different sites showed the growth of violet-colored colonies on culture plates, suggestive of *C. violaceum* (unpublished).
data). The seasonality in the availability of substrates used by *C. violaceum* bacteria can be a determinant factor in the spatial distribution of this species in the Negro River. Our findings are in agreement with those of various authors who observed different concentrations of substrates determining the seasonal fluctuation of bacterial populations in aquatic environments (Casamayor et al., 2002; Pearce et al., 2005; Smith, 2007). Further samplings should be made in the Negro River to confirm this hypothesis.

The methodological strategy chosen to guarantee the separation of the bacterioplankton from eukaryotic organisms, consisted of a pre-filtration step. It is possible that this procedure made it impossible to recover any bacteria bound to particles of sediments in suspension (Lemarchand et al., 2006). The abundance of sediments in suspension is a peculiar characteristic of the Solimões River in Amazonas, (Sioli, 1976; Sioli, 1984). In comparative studies of planktonic bacterial communities associated with particles and communities of free bacteria, or correlating them with biotic and abiotic changes, a richness of OTUs and greater genetic diversity were observed in the communities of bacteria associated with particles, when compared to the OTUs of communities of free bacteria (Crump et al., 1999; Lemarchand et al., 2006). These findings point to the possibility that populations of bacterioplankton of the Solimões River were under-sampled due to the type of strategy adopted for obtaining the bacterial biomass.

The application of molecular biology techniques including the strategy of shotgun cloning and amplification, cloning and sequencing of the SSU rRNA gene of total DNA in aquatic environments, has permitted the analyses of the structure of microbial communities with greater precision, and concomitant inventory of the ecological and biotechnological potential of prokaryote populations from complex environments (Tyson et al., 2004; Venter et al., 2004). However, the genetic databanks utilized for the characterization and taxonomic affiliation of the sequences, NCBI (BLAST) and RDP (*classifier*), still appear to be insufficient in the total coverage of the prokaryotic diversity of unexplored natural environments such as Amazonia. The comparison of the sequences recovered from the Solimões and Negro Rivers, with the sequences deposited in databanks showed that 79 OTUs of the Solimões and 69 OTUs of the Negro did not show similarity with OTUs of any genus known to date, and were partially identified through the identification of their closest parents, showing an immense genetic diversity to be revealed in the near future, where new genera and species will be confirmed within the bacterial divisions sampled in the two Amazon rivers.

**ACKNOWLEDGMENTS**

Research supported by CNPq with a CT-Amazonia grant and a fellowship. We are also grateful to Enedina Nogueira de Assunção for technical assistance. Dr. A. Leyva helped with the translation and editing of the manuscript.

**REFERENCES**


