Microbial Diversity of a Brazilian Coastal Region Influenced by an Upwelling System and Anthropogenic Activity

Juliano C. Cury1,4*, Fabio V. Araujo2, Sergio A. Coelho-Souza3, Raquel S. Peixoto1, Joana A. L. Oliveira4, Henrique F. Santos1, Alberto M. R. Dávila4, Alexandre S. Rosado1

1 Laboratory of Molecular Microbial Ecology, UFRJ, Rio de Janeiro, Brazil, 2 Laboratory of Microbiology, DCIEN, FFP/UERJ, Rio de Janeiro, Brazil, 3 INPeTAm/IBCCF, UFRJ, Rio de Janeiro, Brazil, 4 Laboratory of Computational Biology and Systems, IOC, FIOCRUZ, Rio de Janeiro, Brazil

Abstract

Background: Upwelling systems are characterised by an intense primary biomass production in the surface (warmest) water after the outcrop of the bottom (coldest) water, which is rich in nutrients. Although it is known that the microbial assemblage plays an important role in the food chain of marine systems and that the upwelling systems that occur in southwest Brazil drive the complex dynamics of the food chain, little is known about the microbial composition present in this region.

Methodology/Principal Findings: We carried out a molecular survey based on SSU rRNA gene from the three domains of the phylogenetic tree of life present in a tropical upwelling region (Arraial do Cabo, Rio de Janeiro, Brazil). The aim was to analyse the horizontal and vertical variations of the microbial composition in two geographically close areas influenced by anthropogenic activity (sewage disposal/port activity) and upwelling phenomena, respectively. A lower estimated diversity of microorganisms of the three domains of the phylogenetic tree of life was found in the water of the area influenced by anthropogenic activity compared to the area influenced by upwelling phenomena. We observed a heterogenic distribution of the relative abundance of taxonomic groups, especially in the Archaea and Eukarya domains. The bacterial community was dominated by Proteobacteria, Cyanobacteria and Bacteroidetes phyla, whereas the microeukaryotic community was dominated by Metazoa, Fungi, Alveolata and Stramenopile. The estimated archaeal diversity was the lowest of the three domains and was dominated by uncharacterised marine Crenarchaeota that were most closely related to Marine Group I.

Conclusions/Significance: The variety of conditions and the presence of different microbial assemblages indicated that the area of Arraial do Cabo can be used as a model for detailed studies that contemplate the correlation between pollution-indicating parameters and the depletion of microbial diversity in areas close to anthropogenic activity; functional roles and geochemical processes; phylogeny of the uncharacterised diversity; and seasonal variations of the microbial assemblages.

Introduction

The coastal upwelling consists of an upward movement of bottom seawater and depends on the confluence of meteorological factors and continental morphology. Upwelling systems are characterised by an intense primary biomass production in the surface (warmest) water after the outcrop of the bottom (coldest) water, which is rich in nutrients. The Arabian Sea and coasts of Namibia, Chile, Peru and Benguela display areas where upwelling phenomena occur in a more pronounced way [1–4]. Some studies have confirmed the existence of a typical upwelling system on the southeast Brazilian coast and that the Arraial do Cabo region is the location where the main event occurs [5–7]. In this region, the South Atlantic Central Water (SACW) rises to the surface of the ocean in accordance with the east-northeast winds that occur more frequently between spring and summer and also due to the influence of cyclonic meanders [8–11].

Studies have demonstrated the ecological importance of these systems, with nitrogen loss by nitrification [1,12,13] and increased biological productivity caused by the outcropping of the deep waters [14–17] stimulating fishing activities [18].

Although it is known that the microbial assemblage plays an important role in the food chain of marine systems [19] and that the Brazilian upwelling system drives the complex dynamics of the food chain [20], little is known about the microbial composition present in the central area of such upwelling. Moreover, the microbial composition of a region can vary due factors such as ocean water warming [21], water column depth [22], water mass boundaries [23] and the presence of pollutants [24]. Furthermore, there is little information available concerning the microbial
diversity of the Brazilian coast. Culture-independent studies based on SSU rRNA sequences have shown that coastal systems harbour a vast and uncharacterised diversity of microorganisms representing the three domains of the phylogenetic tree of life [25–28]. We carried out a molecular survey with the aim of assessing microbial assemblages and determining how their variation in the marine waters of Arraial do Cabo are influenced by upwelling events [29] and anthropogenic activities [30].

**Materials and Methods**

**Location, sampling and environmental parameters**

The studied site is located in Arraial do Cabo, at the Cabo Frio upwelling region, Rio de Janeiro state, Brazil (Fig. 1). Two sites were chosen for the water sampling: one site is located in Baia dos Anjos and is more influenced by anthropogenic activity and less by upwelling (PO); the other site is located in an oceanic area where upwelling (RE) occur in the absence of the influence of anthropogenic activities. Physical-chemical parameters (dissolved oxygen – DO, pH, water temperature, salinity, turbidity and conductivity) were measured in situ with a multiparameter probe (YSI, Model). Water samples were collected on 29 February 2008, with a 5-L Van Dorn bottle at the sub-surface (0.5 m) (POS and RES) and at the bottom at depths of 20 m (POF) and 50 m (REF) for PO and RE, respectively. Sub-samples of 100 mL were collected for microbiological counts (coliforms and heterotrophic bacteria), placed in sterilised bags and stored at 4°C on wet ice. For nutrient analysis (ammonia, nitrite, nitrate and phosphate), 1 L of each van Dorn sample was placed in a polypropylene bottle and stored at 4°C on wet ice. For molecular analysis, 1 L of water was collected in triplicate (three van Dorn casts for each sub-surface and bottom depth sampling), placed in autoclaved polypropylene bottles and stored at 4°C on wet ice. All samples stored on wet ice were immediately transported under refrigeration to the laboratory for processing. Twenty microlitres of water from each Van Dorn bottle was separated for bacterial activity and flow cytometry analyses. For bacterial enumeration, three 1.7-mL samples of local water were fixed with 2% paraformaldehyde and frozen. Bacterial concentrations varied around 50% between the samples, but all three samples were pooled before dye input.

Volumes of 1.7 mL of sampled water were used for bacterial activity analysis as well. The first step, the addition of 3H-leucine, was performed under field conditions in triplicate and using a control. After incubation, the samples were frozen until protein extraction.

**Bacterial counting**

Total and thermotolerant coliforms were identified only in sub-surface waters using MPN methods. Serial dilutions of the water samples (10; 1; 0.1; 0.01 and 0.001 mL) were inoculated in lauryl sulphate broth and incubated for 24 h at 35–37°C to determine the total coliform counts. Positive tubes (presence of bubbles and gas production) were reincubated in tubes containing EC broth and incubated for 48 h at 44.5°C to determine the thermotolerant coliform counts. MPN tables were used to determine the numbers of total and thermotolerant coliforms [31]. The total numbers of heterotrophic bacteria were determined by the spread plate method. Aliquots of 0.1 mL and 10× and 100× dilutions were spread onto plates containing Marine Agar. These plates were incubated at 25±3°C for 48 h, after which the CFUs (colony-forming units) were counted [31].

**Flow cytometry and bacterial activity**

Bacterial concentrations were measured with a CytoSence-CytoBuoy flow cytometry equipped with a solid blue laser providing 20 mV at 488 nm, one side scatter (SWS, 446/500 nm) detector and three others to red (chlorophyll-a) fluorescence (FL1- 669/725 nm); orange/yellow (FL-2, 601/651 nm), and green/yellow (FL-3, 515/585 nm) fluorescence, respectively [32]. A yellow-green 0.92 and 10 μm beads (Fluoresbrite Microparticles, Polysciences Inc.-Warrington, PA) were used as an internal standard and to bacterial counts field determination using length sidescatter and average orange fluorescence parameters. Bacterial pooled samples were incubated for 15–30 minutes in the dark with Syber Green I at a final concentration of 0.5 × 10⁻⁴ of the commercial stock solution [33–35]. Samples were run 4 times for 1 minute at a rate of 2 m.s⁻¹ and the discriminator was set to sidescatter and orange fluorescence.
Bacterial activity was assessed according to Smith & Azam [36]. Triplicate samples (1.7 ml) were incubated for 1 h in 2-ml Eppendorf tubes containing L-[4,5-3H] leucine (Amersham TRK 510; specific activity: 73 Ci mmol⁻¹) at a final concentration of 10 nM. The incubations were performed in situ inside thermal boxes to maintain all samples at the same temperatures. One tube was amended with 90 µl of 100% ice-cold trichloroacetic acid (TCA) as a killed control. After 1 h, the incubation was halted by the addition of TCA, and the tubes were frozen. In the laboratory, samples were processed for 3H-protein extraction by centrifugation, and the isotopic activity was determined using a Packard Tricarb 1600 TR liquid scintillation counter with internal calibration and quench correction. Bacterial carbon production was calculated using a protein/carbon conversion factor of 0.86 [37], as used in others studies conducted in this region [20,38].

**DNA extraction**

Each 1-L sample was filtered through a 0.22-µm diameter cellulose filter membrane (Millipore™) and stored at −80°C. Each membrane was macerated with liquid nitrogen and used for DNA extraction. The metagenomic DNA was extracted using the UltraClean™ Soil DNA kit (MOBIO Laboratories – Carlsbad, California) according to the manufacturer’s instructions. DNA integrity was verified by electrophoresis in a 1% agarose gel in TAE buffer.

**PCR-DGGE analysis**

For PCR-DGGE (denaturing gradient gel electrophoresis) analysis of the bacterial community, the 16S rRNA fragments were amplified by PCR using specific primers (Table 1). The samples were collected in triplicate. DGGE of 40 µl of amplicons of the bacterial 16S rRNA from each sample was carried out using a DCode system (BioRad) at 175 V and 60°C for 40 min and then scanned using a Storm PhosphorImager (Molecular Probes) as a killed control. After 1 h, the incubation was halted by the addition of TCA, and the tubes were frozen. In the laboratory, samples were processed for 3H-protein extraction by centrifugation, and the isotopic activity was determined using a Packard Tricarb 1600 TR liquid scintillation counter with internal calibration and quench correction. Bacterial carbon production was calculated using a protein/carbon conversion factor of 0.86 [37], as used in others studies conducted in this region [20,38].

**PCR-DGGE analysis**

For PCR-DGGE (denaturing gradient gel electrophoresis) analysis of the bacterial community, the 16S rRNA fragments were amplified by PCR using specific primers (Table 1). The samples were collected in triplicate. DGGE of 40 µl of amplicons of the bacterial 16S rRNA from each sample was carried out using a DCode system (BioRad) at 175 V and 60°C for 40 min and then scanned using a Storm PhosphorImager (Molecular Probes) as a killed control. After 1 h, the incubation was halted by the addition of TCA, and the tubes were frozen. In the laboratory, samples were processed for 3H-protein extraction by centrifugation, and the isotopic activity was determined using a Packard Tricarb 1600 TR liquid scintillation counter with internal calibration and quench correction. Bacterial carbon production was calculated using a protein/carbon conversion factor of 0.86 [37], as used in others studies conducted in this region [20,38].

**Table 1. Overview of the primers and PCR conditions used to amplify the partial SSU rRNA gene for marine water samples of Arraial do Cabo region.**

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Sequence (5’-3’)*</th>
<th>Used for</th>
<th>PCR Conditions</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>U968F*</td>
<td>AACGCGAAGACCTTAC</td>
<td>DGGE</td>
<td>3 min initial denaturation at 94°C; 35 cycles of denaturation (1 min at 94°C), annealing (1 min at 55°C), and extension (1 min at 72°C); final extension at 72°C for 10 min</td>
<td>[39]</td>
</tr>
<tr>
<td>L1401R</td>
<td>GCGTGTGTAACAGACC</td>
<td>Library</td>
<td>5 min initial denaturation at 95°C; 30 cycles of denaturation (1 min at 92°C), annealing (1 min at 55°C), and extension (1 min at 72°C); final extension at 72°C for 10 min</td>
<td>[40]</td>
</tr>
<tr>
<td>BAC27F</td>
<td>AGAGGTTGATCMGCTCAG</td>
<td>Library</td>
<td>5 min initial denaturation at 95°C; 30 cycles of denaturation (1 min at 92°C), annealing (1 min at 55°C), and extension (1 min at 72°C); final extension at 72°C for 10 min</td>
<td>[40]</td>
</tr>
<tr>
<td>BAC519R</td>
<td>ATTCGGCGGCGCTGG</td>
<td>Library</td>
<td>5 min initial denaturation at 95°C; 30 cycles of denaturation (1 min at 92°C), annealing (1 min at 55°C), and extension (1 min at 72°C); final extension at 72°C for 6 min</td>
<td>[41,42]</td>
</tr>
<tr>
<td>ARCH21F</td>
<td>TTGCGGTTAGCCGCCGA</td>
<td>Library</td>
<td>5 min initial denaturation at 95°C; 30 cycles of denaturation (30 s at 95°C), annealing (30 s at 53°C), and extension (1 min at 72°C); final extension at 72°C for 6 min</td>
<td>[43]</td>
</tr>
<tr>
<td>ARCH519R</td>
<td>GCCGCCGCGCGCGG</td>
<td>Library</td>
<td>5 min initial denaturation at 95°C; 30 cycles of denaturation (1 min at 95°C), annealing (1 min at 55°C), and extension (1 min at 72°C); final extension at 72°C for 10 min</td>
<td>[43,44]</td>
</tr>
</tbody>
</table>

*For DGGE analysis a GC-clamp (5’-GCCCCCCGGCAGCCGGCCGGGCGGGCAGCCGGCCGGCCGGG-3’) was added to the 5’ end of the forward primers.

doi:10.1371/journal.pone.0016553.t001

**Construction of SSU rRNA gene clone libraries and sequencing**

One sample of each triplicate series of samples (repetition number one) was used for SSU rRNA library construction and sequencing. Table 1 shows the name, sequence, PCR conditions and reference for the primers used. For the 16S rRNA from Archaea, nested PCR was performed using the primer pairs ARCH21F-ARCH519R and ARCH21F-ARCH519R for the first and second amplification, respectively. The amplified fragments were ~310, ~530 and ~450 base pairs in length for the Archaea, Eukarya and Bacteria domains, respectively. Agarose gel electrophoresis of 150 µl of each PCR product was performed prior to purification with the QIAquick Gel Extraction Kit (QiaGen) according to the manufacturer’s instructions. Purified amplicons were ligated into the pGEM® T Easy Vector plasmid (Promega). The ligated products were transformed into DH5-α Escherichia coli competent cells. Positive clones were grown in LB medium, and the plasmids were isolated using a miniprep alkaline lysis method [45]. Each insert was sequenced using the BigDye terminator system and an ABI-3730 automatic capillary sequencer (Applied Biosystems).

**Sequence analysis**

The electropherogram sequencing files were processed using the Phred program [46] for base calling and for trimming of vector and low quality (<20) sequences. The high quality sequences located between the rRNA primers were used for further analysis. The prokaryotic sequences were chimera-checked using the
The influence of anthropogenic activities was apparent in POS and POF samples as showed by the thermotolerant coliforms results (Table 2). The total cultivable heterotrophic bacterial counts revealed differences between surface and bottom samples at both sites, with higher values determined for POS and POF compared to RES and REF (Table 2). The bacterial abundance measured by flow cytometry also demonstrated higher values for POS and POF compared to RES and REF, with no observed difference between depths (Table 2). Bacterial activity measured by \(^{15}N\)-leucine incorporation was higher in superficial waters (POS and RES). Considering the superficial layer, in the sample affected by anthropogenic activity (POS), the bacterial activity was 3.3-fold higher than that in the site influenced by upwelling (RES).

### Microbial diversity and taxonomic composition

The estimated values for the OTU richness, diversity index, sample coverage and rarefaction curves of Bacteria, Archaea and Eukarya domains for the four samples analysed are presented in Table 3 and Fig. 2. For all domains, the estimated richness of the phylogenotypes tended to be lower in the PO samples. The Shannon diversity index was statistically lower in the POS sample than in the other three samples for Bacteria and Eukarya domains (Table 3).

Unlike the other two domains, the Archaea presented a significantly higher Shannon diversity index associated with the POS sample compared with the other samples (Table 3). The archaeal domain presented the lowest estimated richness and diversity, calculated using DOTUR\(_{0.05}\), when compared with the other two domains (Table 3 and Fig. 2). The estimated sample coverage for the Archaea ranged from 97% to 99% (Table 3), suggesting that the majority of the phylogenotypes of that domain was represented in the generated libraries. The estimated archaeal richness ranged from 16 (POF) to 25 (REF). Despite the lack of statistical differences in the estimated archaeal OTU richness between the samples, the values of the estimated Shannon diversity index presented some significant differences (Table 3). Top water samples presented higher archaeal diversity compared to bottom water samples for each region, and the POS sample contained a higher diversity compared to the other samples. We carried out a comparison using the Fast UniFrac P Test to define the phylogenetic composition similarities between the samples. The results presented in Table 4 (above values) indicate no differences between POS and POF samples (despite the difference in diversity presented in Table 3), whereas all other comparisons presented highly significant differences. These results indicated a possible heterogeneous distribution of the phylogenotypes between the samples. To investigate the distribution of the phylogenotypes between the samples based on the taxonomic information, we conducted analyses of classification and phylogeny. According to the RDP-Classifier, the archaeal community of the study area was dominated overwhelmingly by members of the Euryarchaeota phylum because almost all archaeal sequences (96.4%) were classified as Euryarchaeota (Fig. 3a). Surprisingly, the RDP-Classifier was unable to classify the majority of these sequences (99.1%) to the class level, even when a threshold of 50% was used (Fig. 3b). Next, we performed a phylogenetic analysis using a representative sequence of each OTU (total of 54 OTUs) defined by a DOTUR\(_{0.05}\) analysis containing all archaeal sequences (1157). After aligning representatives of each OTU using the Aligner tool provided by the SILVA database [61], we found that only uncultivated sequences were recovered as nearest neighbours.

Unlike the RDP-Classifier classification (which classified most of the sequences as Euryarchaeota), most of the nearest neighbours were defined as crenarchaeota-like (uncultivated) in GenBank. We
Table 2: Physical, chemical and biological characteristics of the marine water samples of Arraial do Cabo region used in the molecular analysis.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Sal. (s)</th>
<th>pH</th>
<th>DO (mg L⁻¹)</th>
<th>Red. Oxid.</th>
<th>O₂ (mg L⁻¹)</th>
<th>Nitrate (µM)</th>
<th>Nitrite (µM)</th>
<th>Phosphate (µM)</th>
<th>PO₄ (µM)</th>
<th>Total Nitrogen (µM)</th>
<th>NH₄ (µM)</th>
<th>Bacterial Production (part. mL⁻¹)</th>
<th>Total Coliforms (CFU.100 mL⁻¹)</th>
<th>Thermotolerant Coliforms (CFU.100 mL⁻¹)</th>
<th>Total Heterot. Bacteria* (part. mL⁻¹)</th>
<th>Total Bacterial Production (µg C L⁻¹ h⁻¹)</th>
<th>Total Bacterial Production (part. mL⁻¹)</th>
<th>Total Bacterial Production (µg C L⁻¹ h⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>POS</td>
<td>36.0</td>
<td>8.0</td>
<td>2.9</td>
<td>N.D.</td>
<td>4.25</td>
<td>1.510⁻¹</td>
<td>N.D.</td>
<td>N.D.</td>
<td>N.D.</td>
<td>2.32</td>
<td>4.41</td>
<td>0.28</td>
<td>19,200</td>
<td>428,000</td>
<td>4,41</td>
<td>2.32</td>
<td>4.41</td>
<td>19,200</td>
</tr>
<tr>
<td>POF</td>
<td>36.0</td>
<td>8.0</td>
<td>2.9</td>
<td>N.D.</td>
<td>4.25</td>
<td>1.510⁻¹</td>
<td>N.D.</td>
<td>N.D.</td>
<td>N.D.</td>
<td>2.32</td>
<td>4.41</td>
<td>0.28</td>
<td>19,200</td>
<td>428,000</td>
<td>4,41</td>
<td>2.32</td>
<td>4.41</td>
<td>19,200</td>
</tr>
<tr>
<td>RES</td>
<td>36.0</td>
<td>8.0</td>
<td>2.9</td>
<td>N.D.</td>
<td>4.25</td>
<td>1.510⁻¹</td>
<td>N.D.</td>
<td>N.D.</td>
<td>N.D.</td>
<td>2.32</td>
<td>4.41</td>
<td>0.28</td>
<td>19,200</td>
<td>428,000</td>
<td>4,41</td>
<td>2.32</td>
<td>4.41</td>
<td>19,200</td>
</tr>
<tr>
<td>REF</td>
<td>36.0</td>
<td>8.0</td>
<td>2.9</td>
<td>N.D.</td>
<td>4.25</td>
<td>1.510⁻¹</td>
<td>N.D.</td>
<td>N.D.</td>
<td>N.D.</td>
<td>2.32</td>
<td>4.41</td>
<td>0.28</td>
<td>19,200</td>
<td>428,000</td>
<td>4,41</td>
<td>2.32</td>
<td>4.41</td>
<td>19,200</td>
</tr>
</tbody>
</table>


doi:10.1371/journal.pone.0016553.t002

Although previous studies have demonstrated that cloning a representative sequence of each OTU defined by DOTUR₀.03 as input data. Characterisation of the most similar sequences and the most similar sequences of a specific 18S rRNA primers. Table S1 shows the output occurrences of the most similar sequences and the most similar sequences of a characterised organism deposited in GenBank using a representative sequence of each OTU defined by DOTUR₀.03 as input data. Although previous studies have demonstrated that cloning ap-


<table>
<thead>
<tr>
<th>Library</th>
<th>NSa</th>
<th>OTUsb</th>
<th>Estimated OTU richness</th>
<th>Shannon</th>
<th>ACE</th>
<th>Chao1</th>
<th>ESCc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacteria</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POS</td>
<td>350</td>
<td>58</td>
<td>163 (108; 278)</td>
<td>0.89</td>
<td>2.39</td>
<td>2.20</td>
<td>2.57</td>
</tr>
<tr>
<td>POF</td>
<td>252</td>
<td>96</td>
<td>255 (179; 400)</td>
<td>0.74</td>
<td>3.79</td>
<td>3.61</td>
<td>3.97</td>
</tr>
<tr>
<td>POSc</td>
<td>602</td>
<td>140</td>
<td>452 (329; 655)</td>
<td>0.98</td>
<td>3.41</td>
<td>3.26</td>
<td>3.56</td>
</tr>
<tr>
<td>RES</td>
<td>275</td>
<td>108</td>
<td>217 (171; 296)</td>
<td>0.78</td>
<td>4.19</td>
<td>4.05</td>
<td>4.32</td>
</tr>
<tr>
<td>REF</td>
<td>261</td>
<td>93</td>
<td>238 (175; 351)</td>
<td>0.77</td>
<td>4.02</td>
<td>3.87</td>
<td>4.16</td>
</tr>
<tr>
<td>REc</td>
<td>536</td>
<td>171</td>
<td>385 (308; 505)</td>
<td>0.82</td>
<td>4.47</td>
<td>4.37</td>
<td>4.58</td>
</tr>
<tr>
<td>Eukarya</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POS</td>
<td>317</td>
<td>37</td>
<td>59 (45; 94)</td>
<td>0.95</td>
<td>2.30</td>
<td>2.14</td>
<td>2.46</td>
</tr>
<tr>
<td>POF</td>
<td>341</td>
<td>71</td>
<td>105 (87; 144)</td>
<td>0.91</td>
<td>3.06</td>
<td>2.88</td>
<td>3.24</td>
</tr>
<tr>
<td>POSc</td>
<td>658</td>
<td>98</td>
<td>140 (120; 178)</td>
<td>0.85</td>
<td>3.24</td>
<td>3.12</td>
<td>3.37</td>
</tr>
<tr>
<td>RES</td>
<td>322</td>
<td>81</td>
<td>162 (124; 235)</td>
<td>0.86</td>
<td>3.40</td>
<td>3.23</td>
<td>3.56</td>
</tr>
<tr>
<td>REF</td>
<td>248</td>
<td>81</td>
<td>184 (136; 275)</td>
<td>0.80</td>
<td>3.15</td>
<td>2.91</td>
<td>3.40</td>
</tr>
<tr>
<td>REc</td>
<td>568</td>
<td>142</td>
<td>259 (213; 339)</td>
<td>0.87</td>
<td>3.70</td>
<td>3.54</td>
<td>3.85</td>
</tr>
<tr>
<td>Archaea</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POS</td>
<td>366</td>
<td>16</td>
<td>19 (17; 25)</td>
<td>0.99</td>
<td>2.19</td>
<td>2.10</td>
<td>2.28</td>
</tr>
<tr>
<td>POF</td>
<td>265</td>
<td>10</td>
<td>16 (11; 45)</td>
<td>0.99</td>
<td>1.21</td>
<td>1.08</td>
<td>1.34</td>
</tr>
<tr>
<td>POSc</td>
<td>631</td>
<td>24</td>
<td>36 (27; 72)</td>
<td>0.99</td>
<td>2.21</td>
<td>2.12</td>
<td>2.29</td>
</tr>
<tr>
<td>RES</td>
<td>251</td>
<td>22</td>
<td>23 (22; 27)</td>
<td>0.99</td>
<td>1.72</td>
<td>1.52</td>
<td>1.92</td>
</tr>
<tr>
<td>REF</td>
<td>275</td>
<td>20</td>
<td>25 (21; 41)</td>
<td>0.97</td>
<td>1.42</td>
<td>1.05</td>
<td>1.34</td>
</tr>
<tr>
<td>REc</td>
<td>526</td>
<td>34</td>
<td>50 (40; 82)</td>
<td>0.97</td>
<td>1.57</td>
<td>1.41</td>
<td>1.72</td>
</tr>
</tbody>
</table>

*Number of sequences for each library.

bCalculated with DOTUR at the 3% distance level.

cEstimated sample coverage: Cx = 1-(Nx/n), where Nx is the number of unique sequences and n is the total number of sequences.

Table 3. Estimated OTU richness, diversity indices and estimated sample coverage for SSU rRNA libraries calculated with DOTUR0.03.

...phyla (Fig. 5b). The Arthropoda presented a high relative abundance compared with the phyala Ascomycota, Basidiomycota and Chytridiomycota (Fig. 5c). The Alveolata group presented a greater relative abundance in RES and REF samples compared with POS and POF (Fig. 5a). Most of the sequences classified as Alveolata were members of the class Dinophyceae (dinoflagellates), whereas the other sequences were members of the phylum Apicomplexa and the group Ciliophora (ciliates) (Fig. 5d). The Stramenopiles group demonstrated a higher relative abundance in RES sample compared with POS, POF and REF samples (Fig. 5a). Within Stramenopiles group, members of the phylum Bacillariophyta presented the higher relative abundance in POS, POF and RES samples, whereas members of the order Labyrinthulida showed the higher relative abundance compared to the other phylla in REF sample (Fig. 5c).

**PCR-DGGE profiles**

To obtain an overview of the bacterial and eukaryotic community similarity, we performed DGGE analyses using partial sequences of SSU rRNA. When the 18S rRNA of Bacteria was analysed, the result revealed two main clusters (Fig. 6a): one group of REF samples and the other of the other three sample types. A simple view of the gel indicated that the communities comprising the larger cluster were more similar to each other than to any REF repetition. Within the larger cluster, we observed the formation of two subclusters, one containing the POF repetition cluster and the other containing a POS and a RES repetition. Within the other subcluster, the remaining POS and REP repetitions grouped together. When the 18S rRNA of Eukarya was analysed, the result revealed two main clusters (Fig. 6b): one group of samples of the area influenced by anthropogenic activity (POS and POF) and the other of samples of the area influenced by upwelling (RES and REF). Within each main cluster, repetitions of sampling did not show tendency to group together, indicating a less influence of depth related to the definition of the eukaryotic assemblages (Fig. 6b).

**Discussion**

**Bacterial counting and production**

Bacterial counting and production were higher in the top samples and in the samples influenced by anthropogenic activities (Table 2). Normally, low temperature decreases bacterial production [70,71], and Guenther and colleagues [20] suggested a great variability in the plankton interactions during a downwelling-upwelling cycle. During upwelling events, bacterial production decreased from 0.1 to 0.02 µC.L L−1.h−1, and superficial waters presented higher bacterial production than samples collected at 40 m depth. Carvalho & Gonzalez-Rodriguez [38] simulated mixtures of Arraial do Cabo water types and observed that bacterial production induced plankton primary activity but was benefited after maximum phytoplankton production rates. The dynamics observed in total production, growth and biomass turnover rates should consider the complex shifts in the distribution of activity and physiological states within the community assemblage [72].

**Microbial diversity and assemblages**

We found a low number of phylotypes of Archaea (Table 3) in comparison with previous reports that investigated other environments like soils [73] or sediments [74] using the same approach.
However, richness ranging from 16 to 25 can represent higher diversity in Arraial do Cabo when compared with other marine environments. For example, Galand and colleagues [75] found a richness (estimated by Chao1) that ranged from 3 to 7 phylotypes in arctic marine deep waters (62 to 180 m).

The results confirmed the tendency of crenarchaeotal groups to dominate marine environments (Figs. 3 and S1), especially the marine group I crenarchaeotes (MGI). According to DeLong [76], MGI is a widespread archaeal group that appears to be derived from thermophilic ancestors that invaded diverse non-extreme environments. According to the author, MGI and Group II Euryarchaeota (MGII) are the two most abundant archaeal planktonic groups. Some studies have demonstrated that MGI often represents a substantial fraction of the picoplanktonic community [76,77]. According to Karner and colleagues [78], there are approximately $1.3 \times 10^{28}$ archaeal cells in the world’s oceans, of which ~20% are Crenarchaeota. Despite the evidence for a co-dominance of MGI and MGII in marine environments [76,79], the distribution of OTUs and nearest neighbours in the phylogram (Fig. S1), BLAST searches using the generated clones

Figure 2. Rarefaction curves calculated using DOTUR_{0.03}. It was used the partial sequences of microbial SSU rRNA genes from water samples of Arraial do Cabo region. A, B and C: curves of each library of each domain. D, E and F: curves calculated merging the libraries of the two depths for each point of sample. PO: region influenced by port activity and sewage disposal. RE: open ocean region influenced by upwelling. S: surface. F: bottom (20 m for PO and 50 m for RE).

doi:10.1371/journal.pone.0016553.g002

Figure 3. Taxonomic distribution of archaeal 16S rRNA sequences. Classification was performed using the RDP-Classifier (50% of threshold). PO: region influenced by port activity and sewage disposal. RE: open ocean region influenced by upwelling. S: surface. F: bottom (20 m for PO and 50 m for RE).

doi:10.1371/journal.pone.0016553.g003
Table 4. Statistic significance1 (P values) of differences between prokaryotic communities of water samples of Arraial do Cabo region calculated based on partial sequences of 16S rRNA.

<table>
<thead>
<tr>
<th></th>
<th>POS</th>
<th>POF</th>
<th>RES</th>
<th>REF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Corrected</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Archaea</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POS</td>
<td>-</td>
<td>0.684</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>POF</td>
<td>1</td>
<td>-</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>RES</td>
<td>0.762</td>
<td>0.99</td>
<td>-</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>REF</td>
<td>0.42</td>
<td>1</td>
<td>0.168</td>
<td>-</td>
</tr>
<tr>
<td><strong>Not Corrected</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Archaea</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>POS</td>
<td>-</td>
<td>0.101</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>POF</td>
<td>0.444</td>
<td>-</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>RES</td>
<td>0.128</td>
<td>0.137</td>
<td>-</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>REF</td>
<td>0.071</td>
<td>0.301</td>
<td>0.023</td>
<td>-</td>
</tr>
</tbody>
</table>

1P values of UniFrac P Test are calculated based on 1,000 permutations (pairwise differences). (<0.001) highly significant; (0.001–0.01) significant; (0.01–0.05) marginally significant; (0.05–0.1) suggestive; (>0.1) not significant.
2P values corrected for multiple comparisons using the Bonferroni correction.

and nearest neighbours (data not shown) and the finding that only seven generated sequences formed a clade containing a MGII member (Clade II.C – Fig. S1) led us to believe that MGI is the most abundant archaeon in the waters of the Arraial do Cabo region. Some authors have demonstrated a shift of the MGI/MGII dominance according to the season in which the samples were collected [79,80]. Thus, we must consider the seasonal variance of this MGI dominance.

The anthropogenic impact of sewage disposal near to PO site should be favourable for Crenarchaeota, once this group use ammonia as its sole source of energy [81]. However, the significantly higher abundance of Crenarchaeota than Euryarchaeota was also observed in RE samples. These results suggested that the ammonification process could be associated to both natural (upwelling) and artificial (sewage) eutrophication occurring in the Arraial do Cabo region. According to Kuyers and colleagues [1], 50% of marine N loss is associated to upwelling areas and associated oxygen minimum zones (OMZ). Lam and colleagues [82] showed that MGI plays an important role in the nitrification steps of this loss.

We found an unexpectedly high relative abundance of member of Fungi group in POF, RES and REF samples (Fig. 5a) within the eukaryotic assemblage represented by 18S rRNA sequences. However, most sequences could not be classified in subsequent taxonomic levels (Fig. 5c) because the output of the BLAST search indicated an unclassified fungus (Fungal sp., access number GQ120167) as the sequence related (88% to 100% similarity, Table S1) to these widespread sequences found in the study area. GQ120167 is a complete 18S rRNA gene sequence of an uncharacterised fungus isolated from anoxic water of the Arabian Sea [65]. According to another study by the same group, fungal isolates obtained from anoxic waters of the Arabian Sea can play an important role in the N cycle in this environment, participating in denitrification [83].

Spatial zonation of microbial phylotypes

Some differences related to the distribution of the microbial populations could even be detected between samples that did not display differences related to α-diversity or phylogenetic (Fast UniFrac) analysis.

The differential spatial distribution of phylotypes was clearly apparent when we analysed the dispersal of the archaeal sequences in a phylogram (Fig. S1), in which some clusters were formed exclusively by sequences belonging to a sample collection point. This was the case for clusters VIIb and XI, which contained only RE sequences, and cluster VIII, which contained only PO sequences.

Despite the lack of significant differences between the bacterial community compositions of the samples when we analysed the core data using Fast UniFrac, the bacterial communities presented some specificity in the sequence distribution between samples. This was the case for the Planctomycetes phylum, which was visible only in the upwelling influenced site (Fig. 4a). We discovered a total of twelve clones of the Planctomycetes, eleven of which were found in RE samples. All sequences were classified as members of Planctomycetales order by the RDP-Classifier: one clone was found in the POF sample, three clones in the REF sample and eight clones in the RES sample. These clones were represented by OTUs that formed a separate clade in the phylogram (Fig. S2). Together with archaeon MGI nitrifying organisms [83], groups of Planctomycetales are responsible for the anaerobic oxidation of ammonium (anammox bacteria) in suboxic zones [1,13]. According to Kuyers and colleagues [1], ~1% of the cells presented in the water of the Benguela upwelling system were related to anammox bacteria. The authors found a depletion of ammonium followed by increases in nitrite and nitrate concentrations at depths of 30–80 m. These concentrations are in agreement with the concentrations determined in our study, and a depletion of ammonium followed by increases in nitrite and nitrate concentrations was detected in the REF sample (50 m sample) when compared with the other three samples (Table 2). However, we did not find suboxic conditions in the studied area (Table 2). As showed also by previous studies, we hypothesise that the microorganisms potentially involved in the utilisation of fixed inorganic N (e.g., MGI, Planctomycetales, uncharacterised fungi) are facultative anaerobic organisms.

The Planctomycete sequences obtained in our study were affiliated with members of the genus Rhodopirellula (RDP-Classifier threshold 100% - data not shown). The members of Rhodopirellula are considered aerobic heterotrophic bacteria [84]. However, sequencing of the genome of Rhodopirellula baltica [83] revealed the presence of genes related to fermentation. Furthermore, it is known that some Planctomycete strains are considered micro-aerobic, oligotrophic, heterotrophic microorganisms and that the genus Pirellula includes facultative nitrate reducers [84,86]. One of the nearest neighbours that clustered with the Planctomycetes clade in the phylogram shown in Fig. S2 belonged to a strain of Pirellula sp. According to Woebken and colleagues [13], even under water-oxic conditions, anammox bacteria (e.g., Planctomyce-
tales) can oxidise ammonium when they are associated with the interior of the suspended particles, where the oxygen is depleted by groups of heterotrophic bacteria.

The zonation of Eukarya domain was pronounced to the order Ichthyophonida (296 sequences) that belong to an Incertae sedis of the Metazoa/Fungi main group (Fig. S3). Approximately 93% of these sequences were found in the samples that were influenced by anthropogenic activity, especially in the top sample (~72%), whereas only ~7% of the sequences were found in samples that were influenced by upwelling (Table S1).

Most of these sequences were related to an Ichthyophonida organism isolated from marine invertebrate digestive tracts (access...
no. EU124916) [87] and to *Ichthyophorus hoferi* (EU332789), a well-known cosmopolitan parasite of marine fishes [88]. However, the similarity between these two sequences and our clones was low, ranging from 88% to 95% (Table S1). Our sequences were more similar (93% to 99%) to two other environmental sequences: AY884988 [89], found in a Mid-Atlantic estuary, and AY331730 [90], found in the surface water of the Bay of Fundy (East coast of North America). This finding led us to believe that the generated sequences belonged to uncharacterised microeukaryotes.

Natural and anthropogenic impact

The distribution of some bacterial groups between REF and the other sites was clearly in the present studied area (Fig. 6a; Table 4). Traditionally, depth is accepted as the most important factor defining the zonation of prokaryotic occurrence in marine water [42,77,78,82,91–94]. It is well established that vertical zonations can be determined by gradients of light intensity, temperature, hydrostatic pressure, nutrient availability, oxygen content, among others. Superficial water contains more bacterial activity due to the presence of primary producers, whereas in deep waters, the prokaryotes should demonstrate a greater association with particles. Although most prokaryotes are free-living in oceanic water columns, the detachment of cells associated with sinking particles is an important mechanism that supplies free-living prokaryotic cells in deeper waters [95].

The distribution of microorganisms could also be defined by the characteristics of the different water masses [96–98]. Observing the low temperature reported in REF (table 2), SACW was presented at the deep waters from the upwelling influenced area. The different water conditions could explain the formation of different clusters of archaea-containing phylotypes exclusively from RES/REF (upwelling influenced area) or POS/POF (anthropogenic influenced area) samples (Fig. S1), as well as for the Eukarya domain (Fig. 6b). Moreover, the DGGE dendrogram do not show the formation of subclusters contrasting superficial (S) and deep (F) samples. We concluded that the temperature determined by the water type should be the main regulating factor for microbial assemblages and that the anthropogenic impact decreases its diversity in the Arraial do Cabo upwelling area.

**Supporting Information**

**Figure S1** Phylogram of the archaeal 16S rRNA phylotypes. A representative sequence of each OTU determined by DOTUR0.03 and the nearest neighbors obtained by using of the aligner tool of the Silva database project were used. The sequences of type strains used in the phylogram were obtained from the All-Species Living Tree project (Yarza et al., 2008). The number and the letters after the OTU identification indicate the number of sequences clustered in each OTU and the sample where the sequences were found, respectively. PO – area influenced by anthropogenic activity; RE – open ocean area influenced by the upwelling phenomena; S – superficial sampling; F – bottom sampling (20 m for PO and 50 m for RE). The phylogram was calculated with MEGA 4.0 using neighbor-joining method and Jukes-Cantor model. Numbers at the branches show bootstrap percentages (above 50% only) after 1000 replications of bootstrap sampling. The sequence of *Abiotrophia defectiva* (Bacteria) was set as outgroup.

**Figure S2** Phylogram of the bacterial 16S rRNA phylotypes. A representative sequence of each OTU determined by DOTUR0.03 and the nearest neighbors obtained by using of the aligner tool of the Silva database project were used. The number and the letters after the OTU identification indicate the number of sequences clustered in each OTU and the sample where the sequences were found, respectively. PO – area influenced by anthropogenic activity; RE – open ocean area influenced by the upwelling phenomena; S – superficial sampling; F – bottom sampling (20 m for PO and 50 m for RE). The phylogram was calculated with MEGA 4.0 using neighbor-joining method and Jukes-Cantor model. Numbers at the branches show bootstrap percentages (above 50% only) after 1000 replications of bootstrap sampling. The sequence of *Caldisphaera draconis* (Archaea) was set as outgroup.

**Figure S3** Phylogram of the 18S rRNA phylotypes. A representative sequence of each OTU determined by DOTUR0.03 and the nearest neighbors obtained by using of the aligner tool of the Silva database project were used. The number and the letters after the OTU identification indicate the number of sequences clustered in each OTU and the sample where the sequences were found, respectively. PO – area influenced by anthropogenic activity; RE – open ocean area influenced by the upwelling phenomena; S – superficial sampling; F – bottom sampling (20 m for PO and 50 m for RE). The phylogram was calculated with
MEGA 4.0 using neighbor-joining method and Jukes-Cantor

Acknowledgments

Table S1 Closest relative sequences obtained with NCBI-Blast search using generated partial 18S rRNA

References

Table S1 Closest relative sequences obtained with NCBI-Blast search using generated partial 18S rRNA sequences from water samples of Arraial do Cabo region. PO: region influenced by port activity and sewage disposal. RE: open ocean region. S: surface. F: bottom.


Table S1 Closest relative sequences obtained with NCBI-Blast search using generated partial 18S rRNA sequences from water samples of Arraial do Cabo region. PO: region influenced by port activity and sewage disposal. RE: open ocean region. S: surface. F: bottom.


Table S1 Closest relative sequences obtained with NCBI-Blast search using generated partial 18S rRNA sequences from water samples of Arraial do Cabo region. PO: region influenced by port activity and sewage disposal. RE: open ocean region. S: surface. F: bottom.


