Biogeography of actinobacterial communities in hot springs

Angel Valverde, Marla Tuffin, Don A. Cowan*

Institute for Microbial Biotechnology and Metagenomics, University of the Western Cape, Bellville 7535, Cape Town, South Africa.

*Corresponding author: Don A. Cowan. dcowan@uwc.ac.za
Ph: +27 (0)21 959 2083; Fax: +27 (0)21 959 3505

Abstract

Actinobacteria are ubiquitous in soil, freshwater and marine ecosystems. Although various studies have focused on the microbial ecology of this phylum, data is scant on the ecology of actinobacteria endemic to hot springs. Here, we have investigated the molecular diversity of eubacteria, with specific focus on the actinobacteria in hot springs in Zambia, China, New Zealand and Kenya. Temperature and pH values at sampling sites ranged between 44.5-86.5°C and 5-10, respectively. Non-metric multidimensional scaling analysis of 16S rRNA gene T-RFLP patterns showed that samples could be separated by geographical location. Multivariate analysis showed that actinobacterial community composition was best predicted by changes in pH and temperature, whereas temperature alone was the most important variable explaining differences in bacterial community structure. Using 16S rRNA gene libraries, 28 major actinobacterial OTUs were found. Both molecular techniques indicated that many of the actinobacterial phylotypes were unique and exclusive to the respective sample. Collectively, these results support the view that both actinobacterial diversity and endemism are high in hot spring ecosystems.

Keywords: Actinobacteria, biogeography, diversity, hot springs

INTRODUCTION

Actinobacteria are ubiquitous highly diverse microorganisms [2,18,26] involved in the turnover of organic matter. Although actinobacteria are often considered as soil-borne
bacteria, these organisms are also known as pathogens of animals, including humans, and plants [29,34,52]. Other actinobacterial species form beneficial associations with plants [4] and insects [22]. Furthermore, actinobacteria are well-known producers of a vast array of secondary metabolites [55], many of which have useful applications in medicine, veterinary and agriculture. It therefore follows that an understanding of actinobacterial distribution in the environment is important in deciphering the ecological role of these organisms and for biotechnological bioprospecting.

Hot springs are often isolated habitats, ideal for studies of the interactions between organisms and for their ability to adapt to extreme conditions. Physical isolation may create an “island effect”, leading to evolutionary adaptations resulting in highly divergent structural and functional characteristics [38]. Hot spring communities are also considered as analogues of early earth and even extraterrestrial conditions [7]. Moreover, thermophiles and their products are of considerable biotechnological interest, having found effective applications in fields as diverse as PCR and biohydrometallurgy [16].

Molecular approaches have revolutionized microbial ecology and revealed that bacteria, like animals and plants, exhibit biogeographic patterns shaped by distance and historical and contemporary environmental conditions (e.g. [19,42]).

Microbial biography studies in hot springs, a great number carried out in Yellowstone National Park, have shown these environments containing highly distinct community structures, where archaea, cyanobacteria, chloroflexi and acidobacteria are among the microorganisms most commonly found [35,40,50-51,23,56]. Such studies have typically focused on the broader microbial community and very little is known on the biogeography of actinobacteria in hot springs. For instance, Walker et al. [53] found several sequences belonging to Mycobacterium sp. in pore water extracted from rocks collected in Norris Geyser Basin. The, study of two alkaline-silica hot springs by Miller et al. [35] found that only 3 of 391 bacterial operational taxonomic units (OTUs) could be classified as Actinobacteria. In addition, to our knowledge, only the work by Song et al. [49] has specifically addressed the diversity of Actinobacteria in hot springs. However, the actinobacterial distribution in these environments has not been related to physicochemical properties other than temperature. Clearly, more research is needed in order to elucidate the biogeographic patterns and environmental factors that shape the structure of this important subgroup of bacteria in hot springs. Determining factors that control actinobacterial community structure will play an important role in identifying
sampling sites for bioprospecting. Here we provide additional insight into the structure of actinobacterial communities in hot springs, with a comparison of four thermal sites, widely separated geographically and by physicochemical properties. Since subgroups can be controlled by factors other than those that influence the entire community, the biogeographic pattern of the total bacterial community in these sites was also investigated.

**MATERIALS AND METHODS**

**Site description and sampling**

Four different locations with contrasting environmental conditions (Table 1) were chosen for the analysis. Kenyan sediments were taken from Lake Magadi hot springs (S01° 53.219’, E36° 17.536’), located in Great Rift Valley, Kenya. The area is an arid tropical zone where tectonic activity has created a series of shallow depressions. Surface evaporation rates exceed the rate of inflow of water allowing the dissolved minerals to concentrate into a caustic alkaline brine with CO$_3^{2-}$ and Cl$^-$ as major anions, creating a pH of 8.5 to >12.

Zambian sediments were collected from Gwisho hot spring (S15° 59.526’, E27° 14.520’), located in Lochinvar National Park. The Gwisho hot springs occur along a geological fault where the water rises by convection with temperatures ranging from 60° to 90° C. There are high concentrations of sodium, chlorine, calcium and sulphates in the water.

Chinese hydrothermal sediments were sampled from Rehai thermal area (N24° 51.213’, E98° 23.456’), near Tengchong, Yunnan Province. The Rehai geothermal field is located in Tengchong County, in the western Yunnan Province. It extends over 10 km$^2$ where alkaline boiling springs, spouting boiling springs, intermittent spouting springs, hydrothermal eruption vents, high-temperature fumaroles, and widespread areas of steaming ground are commonly found.

New Zealand samples were collected from Tokaanu geothermal site (S38° 58.249’, E175° 46.114’), North Island, New Zealand. Tokaanu is one of the major geothermal fields of the Taupo volcanic zone (TVZ), situated in the North island of New Zealand. The TVZ covers an area approximately 30 kilometers wide by 150 kilometers long and contains 23 known geothermal fields consisting of a number of geysers, sinters deposits, hot springs and pools, steaming cliffs, fumaroles, steam vents and seepages.
Three sediment samples were collected at each site, introduced into sterile 50-mL Falcon conical tubes, and stored immediately on dry ice. Water pH was measured in situ using pH strips (Merck) and temperature was measured using a Solomat 52°C temperature monitor. Anion analysis was performed using ion chromatography by the Environmental Laboratory at Stellenbosch University (South Africa), following standard procedures.

DNA extraction
Metagenomic DNA was extracted from sediments using the protocol described by Ellis et al. [12] with minor modifications. Briefly, sediments (250 mg) were incubated with 1 ml of buffer (100 mM Tris-HCl, 100 mM EDTA [pH 8.0], 100 mM phosphate buffer [pH 8.0], 1.5 M NaCl, 1% CTAB), 10 μl proteinase K (20 mg/ml) and 1 mg of lysozyme at 37°C for 2 hours with agitation (225 rpm). Subsequently, 120 μl of 20% SDS was added, and the samples were incubated at 65°C for 1 hour. The samples were then centrifuged at 5,000 xg for 10 min. The supernatant was transferred to a new tube and consecutively extracted with an equal volume of phenol, phenol-chloroform-isoamyl alcohol (25:24:1) and chloroform-isoamyl alcohol (24:1). The aqueous phase was precipitated with 0.6 volume of isopropanol. After centrifugation, pellets were washed with 300 μl of 70% ethanol and air dried before resuspending in 50 μl of ultrapure water. Genomic DNA was extracted in triplicate, pooled and purified using PVPP spin columns [6].

Terminal-Restriction Fragment Length Polymorphism (T-RFLP) Analysis
Bacterial 16S rRNA gene amplification was performed using primer pair 341F (5’-CCTACGGAGGCAGCAI-3’) [21], and 908R (5’-CCGTCAATTCTMTTTGAGTTI-3’) [25]. For actinobacterial-specific amplifications, primers 226-243-F (5’-GGATGAGCCCGCGCCCTA-3’) [17] and (A3R) 1414-1430-R (5’-CCAGCCCCACCTTCGAC-3) [36] were used. Bacterial primers were modified by the addition of inosine at the 3’ end in an attempt to broaden their target scope [3]. In both primer pairs, the forward primer was labeled with 6’ carboxyfluorescein (6-FAM).

Each PCR mixture (50 μl) contained 2U of DreamTaq (Fermentas, Vilnius, Lithuania), 0.5 μM each primer, 0.1 mM dNTPs, 10 μg of BSA, and 20 ng of DNA. Cycling conditions for bacteria were: 4 min at 94°C; 25 cycles of 1 min at 94°C, 1 min at 55°C and 2 min at 72°C; and a final extension of 10 min at 72°C. For actinobacteria: cycling
conditions were 4 min at 94°C; 25 cycles of 30 s at 94°C, 2 min at 68°C and 60 s at 72°C; 10 min at 72°C. PCR products were combined from three amplification reactions per sample, verified by agarose gel electrophoresis and purified with NucleoSpin Extract II (BD Biosciences Clontech, Japan). Approximately 200 ng of purified products were digested in separate reactions using TaqI, AluI and HhaI restriction enzymes (Fermentas). After purification as above, samples were subjected to capillary electrophoresis using the Applied Biosystems DNA Sequencer 3130 (Applied Biosystems, Foster City, California, USA). Terminal restriction fragments (T-RFs) data generated by Peak Scanner software v1.0 (Applied Biosystems) were filtered and binned by the method developed by Abdo et al. [1]. For each sample, peaks over a threshold of 50 fluorescence units were used and T-RFs of < 30 bp were excluded from the analysis to avoid detection of primers.

**Statistical analysis of T-RFLP fingerprints**

Non-parametric multivariate statistical analysis was performed using the PRIMER 6 software package (PRIMER-E Ltd, Plymouth, UK). Sample-similarity matrices were generated using Sorensen coefficient on presence/absence data [9]. The community structure was explored by ordination using non-metric multidimensional scaling (NMDS). An analysis of similarity (ANOSIM), performed on the resemblance matrix, was used to test for differences in bacterial community structure between predefined groups [8]. The significance of the Anosim R-statistic was tested by Monte Carlo randomization with 999 permutations. To control for false-positive errors that occur during multiple statistical comparisons, P-values were subjected to the Bonferroni correction [41]. The influence of temperature and pH on the community structure of the samples was assessed using the BEST analysis [10]. The significance of the correlations between two distance matrices was calculated using the Mantel test [27] with 999 matrix permutations. Mantel test and other statistical analyses were performed using R software (http://www.r-project.org/).

**16S rRNA gene clone library construction and phylogenetic analysis**

Representative clone libraries were constructed after pooling equal amounts of amplicons from the individual samples for each location obtained with 226-243-F and 1414-1430-R primers. Aliquots of the pooled products were cloned into *Escherichia*
coli DH5α using pGEM-T cloning kit (Promega, Madison, Wisconsin, USA) and transformants were selected by blue-white screening. The presence of correctly sized inserts was confirmed by colony PCR (using M13F/R primers). ARDRA analysis (using AluI and Rsal) was used to dereplicate clones. Restriction patterns were visualized on 2% agarose gels and analysed using Gelcompare II (Applied Maths, Keistraat, Belgium).

Plasmid DNA was extracted using QIAprep Spin Miniprep kit (Qiagen GmbH, Hilden, Germany) and a representative of each unique restriction pattern sequenced using the vector primer M13F with an ABI 3130 DNA Sequencer (Applied Biosystems).

Putative chimeric sequences were filtered using Bellerophon [20]. Sequences of > 97% identity were grouped into OTUs using MOTHUR [47]. Taxonomic assignments of representative OTUs were obtained using the Classifier tool [54] (confidence threshold 80%) from the Ribosomal Database Project II (http://rdp.cme.msu.edu/) [11]. Sequences were aligned using the SINA aligner from the SILVA project (http://www.arb-silva.de) [39] and imported into ARB [31] where alignments were manually curated. The phylogenetic tree was constructed by the neighbour-joining method [43] after the Jukes-Cantor model for nucleotide substitution was applied. The robustness of the tree topology was evaluated by bootstrap analysis [13] based on 1000 resamplings. Sequences obtained in this study were deposited in the NCBI GenBank database under accession numbers JN806355-JN806382.

Diversity analyses were performed on aligned DNA sequences of 671-884 nucleotides in length. Intracommunity diversity was calculated using rarefaction, ACE and the Shannon-Weaver diversity index implemented in MOTHUR. All analyses were performed using the furthest neighbor algorithm with a 3% cut-off (species level) [46]. Clones that were not sequenced but shared the same RFLP pattern were duplicated in alignments.

Intercommunity OTU composition was evaluated with UniFrac analysis (http://bmf2.colorado.edu/unifrac/) [30]. Jackknife analysis was used to assess the confidence of tree nodes. Clone libraries were compared using J-Libshuff analysis [46].

RESULTS

Using molecular analysis of the 16S rRNA gene, the eubacterial diversity, with specific focus on the actinobacterial composition, was compared between four geographically separated thermal sites in Zambia (ZA), China (CH), New Zealand (NZ) and Kenya.
T-RFLP community structure analysis

The 16S rRNA gene amplicons were digested with *Taq*I, *Alu*I and *Hha*I. Mantel test analysis revealed a significant pairwise correlation between T-RFs matrices at all possible combinations for the three restriction enzymes (Mantel R between 0.379 and 0.901 for actinobacteria, and between 0.508 and 0.828 for bacteria, all P=0.001). Thus, all data presented here refer to restriction enzyme *Taq*I, for both bacteria and actinobacteria.

Analysis of actinobacterial T-RFs data for all samples identified a total of 56 distinct T-RFs ranging in size from 33 to 941 bp. The highest number of T-RFs was observed in KE-3 (17) and the lowest in CH-3, NZ-2 and NZ-3 (8) (Table 1). Analysis of bacterial T-RFs data identified a total of 84 distinct fragments that ranged in size from 58 to 583 bp. The highest number of bacterial T-RFs was observed in ZA-1 (40), and the lowest in KE-2 and KE-3 (22) (Table 1). A Venn diagram showing the distribution of the TRFs between the sites is presented in supplementary figure (Online Resource 1). For actinobacteria, a significant negative correlation between the number of T-RFs and temperature was found (Spearman’s $\rho=-0.75$, P=0.005). Differences in the T-RFLP patterns between all samples were visualized in a NMDS plot (Figure 1). Both bacteria and actinobacteria showed a clear separation of samples according to their geographic origin. Analysis of similarity (ANOSIM) showed that the community composition of both actinobacteria and bacteria differed between sampling sites (ANOSIM R=0.642, P=0.021 and R=0.875, P=0.001; respectively). However, all pairwise comparisons between sites were not significant after Bonferroni correction.

Temperature and pH were negatively correlated ($\rho=-0.65$, P=0.024) and strongly shaped the structure of the microbial communities as showed by BEST analysis. Thus, for the actinobacteria the highest Spearman’s rho ($\rho$) correlation of 0.362 (P=0.014) was due to a combination of temperature and pH, whereas for the bacterial data the highest correlation ($\rho=0.709$, P=0.001) was due to temperature alone.

Mantel test confirmed that geographic distance was also a significant factor influencing both bacterial and actinobacterial composition (Mantel R=0.246 and R=0.145, respectively, both P≤0.05).
**Actinobacterial 16S rRNA gene library analysis**

A total of 400 actinobacterial clones were subjected to RFLP analysis, identifying 51 unique banding patterns (data not shown). A total of 28 different OTUs (97% similarity cut-off) were found, of which 19 were affiliated as “actinobacteria” and 9 could not be classified (“unclassified bacteria”).

Phylogenetic analyses showed that a high proportion of actinobacterial sequences were only distantly related to other environmental clones (50-97% similarity values), with very few showing high similarity to cultured isolates (> 97%) (Table 2, Figure 2). Those clones that could be affiliated at the species level (> 97% identity) belonged to the genera *Mycobacterium* (4.6% abundance), *Glycomyces* (1.3%), *Aeromicrobium* (1.3%) and *Couchiplanes* (1.3%), all members of the order Actinomycetales.

Diversity indices showed a difference between the four clone libraries (Table 3), with the libraries from KE and CH being the most diverse and that from NZ the least. The ACE values and the fact that rarefaction curves (Figure 3) reached an asymptote both indicate that in the ZA, KE and NZ libraries, an accurate assessment of actinobacterial diversity was obtained. In contrast, the rarefaction curve and ACE values for the CH library indicates that sampling was below saturation and more sequences would be required to generate a stable estimate of species richness. The Shannon-Weaver index values also corroborate these findings, indicating that the KE and CH libraries were the most diverse. Results from Š-Libshuff analysis of the 16S rRNA gene clone libraries indicated that the communities differed significantly between all pairwise combinations (P≤0.05, after Bonferroni correction). UniFrac analysis of 16S rRNA gene clone libraries supported the conclusion that location had an impact on the composition of the actinobacterial community and that the result was significant. A greater similarity between KE and ZA libraries relative to CH and ZA libraries was observed, forming a separate cluster (Online Resource 2).

**DISCUSSION**

Molecular fingerprinting methods, such as T-RFLP, allow descriptions of broad-scale changes in the taxonomic composition of bacterial communities [28] although the limitations and disadvantages of the technique are well documented (e.g. [48,57]). Our results are in agreement with Zhang et al. [57] and showed that the choice of restriction enzymes in the T-RFLP technique generated highly variable species richness
and diversity indices (data not shown). As a consequence, resemblance matrices based on presence/absence data were used in this study. However, it is important to note that in spite of these differences, the Mantel test analysis revealed consistent patterns in the bacterial and actinobacterial distribution, showing that this technique was suitable for community structure determination in this study [24,57].

In the last decade, information gathered from molecular surveys suggests that microbial biogeographic patterns are shaped by history and environmental factors (reviewed in Martini et al. [32]). For instance, pH has been found to be the best predictor of the continent-scale patterns exhibited by soil bacteria, with the highest diversity in soils with near-neutral pH [14]. Miller et al. [35], analyzing two hot spring bacterial communities along the temperature gradients in YNP, showed that diversity decreased with increasing temperature. Additionally, they found that community similarity decayed exponentially with increasing differences in temperature between samples but not with distance. In contrast, Papke et al. [38] demonstrated that distance effects dominated cyanobacterial diversity in hot springs from North America, Japan, New Zealand and Italy.

Here we show that actinobacterial and bacterial communities were composed of vastly different diversities. The NMDS analysis of T-RFLP profiles showed a clear delineation in bacterial and actinobacterial community structure between samples collected from different locations, with a very pronounced separation between the samples from New Zealand (the hottest site) and the remaining sites. The actinobacterial community composition was best predicted by differences in pH and temperature, whereas temperature alone was the most important factor explaining differences in structure for bacterial communities. Using T-RFLP and clone library analyses, we found that actinobacterial richness and diversity decreased with increasing temperature. In contrast to Miller et al. [35], we did not observe a decrease in the number of bacterial OTUs at the site with the highest temperature, and as previously found [40], the lowest bacterial richness was found at sites with the lowest temperature (Kenya samples). These findings suggest that other physicochemical parameters may play a role in structuring the bacterial community at this site, and support the idea that factors driving variation in bacterial community depend on spatial scale [33]. Nevertheless, as can be inferred from Table 1 and rarefaction curves (Figure 3), and has been shown in other studies [5,45], some discrepancy exists in the number of actinobacterial OTUs observed using both methodologies. Clearly, further research is required in order to elucidate the factors that
shape the structure of bacterial communities at these sites.

Several studies have shown that dispersal limitations exist for thermophiles, leading to the emergence of regionally distinct populations (e.g. [38,40,56]). ANOSIM and Mantel test analyses have shown that geographic distance between the sampling sites was positively correlated with differences in the actinobacterial communities. Moreover, Unifrac analysis, that considers phylogenetic lineages and not only shared OTUs, showed Zambia and Kenya hot springs (less distant sites) having a more similar actinobacterial community structure (Online Resource 2). Taken together, these findings, and those from Song et al. [49], indicate that the differences in actinobacterial community composition in hot springs are a function of both environmental heterogeneity and geographic distance [32,33,42].

Phylogenetic analyses revealed a high diversity of actinobacterial OTUs in hot springs samples, the majority of sequences belonging to unclassified taxa. Only four of twenty-eight OTUs showed high homology (>97%) to known actinobacterial species: *Couchiplanes caeruleus* (OTU 1), *Glycomyces harbinenis* (OTU 2) *Aeromicrobiunm panaciterrae* (OUT 6), and *Mycobacterium parascrofulaceum* (OTU 5). Interestingly, *M. parascrofularaceum* has also been found in acidic (pH 3.0-3.3) hot springs in Yellowstone National Park [44]. The fact that this OTU was sampled in Zambia and Kenya and over a range of pHs (pH 5-10), suggests that this species may be ubiquitous in hot springs. The bulk of the OTUs showed widely varying levels of similarity (50-97%) to environmental sequences recovered from different habitats, including hot springs, radioactive soils, anthracene-contaminated soils and prairie soils (Table 2). As in other environmental studies, several of the sequences identified showed some similarity to non-thermophilic representatives in the databases [35, 49]. However, sequences from “extreme” environments are poorly represented in these databases, and studies of soil bacterial communities have received considerably more attention than those from hot springs.

Interestingly, the majority of those OTUs (7 of 9) that could not be classified as actinobacteria were found in the clone library from China samples (Figure 2). In agreement with these findings, Song et al. [49] also reported some degree of non-specificity of the primers used, as they found several sequences related to formerly candidate phylum OP10 (now described as phylum Armamonadetes). Moreover, they showed highest diversity (15 RFLP types) in the Tengchong sample (12 km from our sampling location). Fewer OTUs were retrieved from the clone library prepared from
our China samples, but rarefaction analysis indicates that the library was not exhaustively sampled. This was further confirmed by the number of TRFs obtained. Collectively, these findings show that this site supports unusually high actinobacterial diversity, and could be a valid target for future bioprospecting. The most striking result is that a very high proportion of the phylotypes detected were unique to the site: i.e., detected in only one sample (Figures 1 and 4). Moreover, the phylotype abundance distribution showed the dominance of a few OTUs in each library (Figure 4), whereas most of the OTUs were relatively rare, resembling the classic “long tail” phenomenon [15]. Only two OTUs were shared between the clone libraries; OTU 5 between Kenyan and Zambian hot springs, and OTU 12 between the Chinese, New Zealand and Zambian sites, and was coincidentally the most abundant in China and New Zealand libraries (39% and 60% of the clones, respectively) (Figure 4). Noticeably, two closest relatives to these OTUs have been found in Yellowstone National Park (*M. parascrofulaceum* and sequence DQ324877, respectively).

The results from this study are in accordance with previous findings [38,56] and suggest that hot springs contain highly endemic populations of microorganisms. We are necessarily extrapolating from small sample sizes and numbers to entire hot spring communities. However, the data presented here is supported by the recent work by Nemergut et al. [37] showing that macro-scale habitats structure bacterial distribution and that there is a positive relationship between the relative abundance of an organism and its distribution across assemblages. Detectable bacteria are confined to single assemblages and the most cosmopolitan taxa are also the most abundant in individual assemblages. We suggest that the emergence of significant and robust patterns in this study indicates that actinobacterial communities are highly diverse and endemic to hot springs.

**Acknowledgements**

The authors gratefully acknowledge the National Research Foundation (South Africa) for support of this research.
References


Acad Sci USA 82:6955-6959


Figure legends

Fig. 1 Non-metric multidimensional scaling based on T-RFLP profiles. A, actinobacteria (stress = 0.13); B, bacteria (stress = 0.09). ZA, Zambia; CH, China; NZ, New Zealand; KE, Kenya

Fig. 2 Neighbor Joining tree of representative 16S rRNA sequences described in this study. Only bootstrap values ≥ 50% are shown. Bar, 0.1 substitutions per nucleotide position

Fig. 3 Rarefaction curves indicating the observed number of OTUs at a genetic distance of 3% in hot springs sediments. Kenya (squares), Zambia (triangles), China (circles), New Zealand (asterisks)

Fig. 4 Rank-abundance plots of OTUs observed in 16S rRNA clone libraries. Only the most abundant and/or shared OTUs (in bold) are shown
Figure 1
Figure 3
Figure 4

China

OTUs rank

Zambia

OTUs rank

New Zealand

OTUs rank

Kenya

OTUs rank

Frequency
Table 1. Physicochemical and T-RFLP data of the hot springs sediments sampled.

<table>
<thead>
<tr>
<th>Sample</th>
<th>pH</th>
<th>Temperature (ºC)</th>
<th>Number of TRFs</th>
<th>Bacteria</th>
<th>Actinobacteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZA-1</td>
<td>5.0</td>
<td>68.3</td>
<td>40</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>ZA-2</td>
<td>5.0</td>
<td>59.3</td>
<td>32</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>ZA-3</td>
<td>6.0</td>
<td>72.7</td>
<td>30</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>CH-1</td>
<td>8.0</td>
<td>66</td>
<td>34</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>CH-2</td>
<td>7.5</td>
<td>58.8</td>
<td>34</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>CH-3</td>
<td>8.5</td>
<td>62.5</td>
<td>32</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>NZ-1</td>
<td>6.5</td>
<td>86.5</td>
<td>31</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>NZ-2</td>
<td>6.5</td>
<td>86.5</td>
<td>31</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>NZ-3</td>
<td>6.5</td>
<td>86.5</td>
<td>34</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>KE-1</td>
<td>9.5</td>
<td>39.4</td>
<td>23</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>KE-2</td>
<td>9.5</td>
<td>44.5</td>
<td>22</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>KE-3</td>
<td>10</td>
<td>51.9</td>
<td>22</td>
<td>17</td>
<td></td>
</tr>
</tbody>
</table>

*aSampling sites. Zambia (ZA), China (CH), New Zealand (NZ), Kenya (KE)*
Table 2. 16S rRNA clones identified in hot springs sediments.

<table>
<thead>
<tr>
<th>OTU</th>
<th>Representative sequence</th>
<th>Accession no.</th>
<th>Closest sequence match with BLASTN, origin (accession number)</th>
<th>Similarity (%)</th>
<th>Closest type strain</th>
<th>Similarity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>C1_KE</td>
<td>Couchiplanes caeruleus, soil (NR_026295)</td>
<td>99 Couchiplanes caeruleus (X93202)</td>
<td>99.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>C2_KE</td>
<td>Glycomyces harbinensis, soil (AJ293747)</td>
<td>98 Glycomyces harbinensis (D85483)</td>
<td>98.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>C3_KE</td>
<td>Uncultured bacterium, soil (FJ478553)</td>
<td>96 Streptomyces albicatios (AY999901)</td>
<td>87.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>C4_KE</td>
<td>Uncultured bacterium, lead-zinc mine (EF612367)</td>
<td>98 Frankia alni (CT573213)</td>
<td>88.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>C6_KE</td>
<td>Mycobacterium sp., tap water (GU084181)</td>
<td>99 Mycobacterium parascrofulaceum (ADNV01000350)</td>
<td>99.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>C7_KE</td>
<td>Uncultured bacterium, Hanford 300 area (HM186589)</td>
<td>95 Thermosiphon atlanticus (AJ577471)</td>
<td>51.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>C8_KE</td>
<td>Aeromicrobium sp., soil (JF824798)</td>
<td>99 Aeromicrobium panactrierae (AB245387)</td>
<td>97.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>C10_KE</td>
<td>Uncultured bacterium, air dust (EF683032)</td>
<td>98 Nocardoides iriomotensis (AB544079)</td>
<td>96.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>C11_KE</td>
<td>Uncultured Modestobacter, Taklamatan desert (JF411374)</td>
<td>96 Modestobacter versicolor (AJ871304)</td>
<td>96.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>C12_KE</td>
<td>Uncultured bacterium, soil contaminated with anthracene (HM438001)</td>
<td>97 Aciditerrimonas ferrireducens (AB517669)</td>
<td>88.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>C14_NZ</td>
<td>Uncultured bacterium, saline-alkaline soil (JN037894)</td>
<td>96 Streptomyces glaucomsporus (AB184664)</td>
<td>89.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>C15_NZ</td>
<td>Uncultured bacterium, geothermal system, Yellowstone National Park (DQ324877)</td>
<td>99 Nitrococcus oceanis (AY033493)</td>
<td>87.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>C21_CH</td>
<td>Uncultured bacterium, Hanford 300 area (HM186978)</td>
<td>95 Laceyella putea (AF1385736)</td>
<td>49.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>C23_CH</td>
<td>Uncultured actinobacterial, sludge (CU924270)</td>
<td>97 Kribella yunnanensis (CP000127)</td>
<td>49.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>C28.CH</td>
<td>Uncultured bacterium, radioactive-waste site (GQ263231)</td>
<td>95 Thermoleophilum minutum (AJ458464)</td>
<td>88.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>C25.CH</td>
<td>Uncultured bacterium, prairie soil (EU132603)</td>
<td>96 Streptomyces panacagri (AB245388)</td>
<td>89.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>C29.CH</td>
<td>Uncultured bacterium, soil (EU132572)</td>
<td>90 Frankia alni (CT573213)</td>
<td>50.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>C30.CH</td>
<td>Uncultured bacterium, hot spring, Thailand (AY555773)</td>
<td>93 Thermoleophilum minutum (AJ458464)</td>
<td>48.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>C32.CH</td>
<td>Uncultured bacterium, praire soil (FJ479127)</td>
<td>96 Thermotoga subterranea (U22664)</td>
<td>48.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>C33.CH</td>
<td>Uncultured bacterium, soil (EF020292)</td>
<td>96 Sporichthya polymorpha (AB025317)</td>
<td>70.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>C37.CH</td>
<td>Uncultured bacterium, rizhosphere soil (EU786138)</td>
<td>98 Acidothermus cellulolyticus (CP0000481)</td>
<td>89.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>C39.ZA</td>
<td>Uncultured bacterium, soil contaminated with anthracene (HM438001)</td>
<td>99 Streptomyces capillospiralis (AB184577)</td>
<td>88.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>C40.ZA</td>
<td>Uncultured bacterium, soil (EF516773)</td>
<td>98 Luedemannella flavus (AB236959)</td>
<td>88.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>C45.ZA</td>
<td>Uncultured bacterium, soil (FJ478871)</td>
<td>96 Streptomyces canigerans (AB184640)</td>
<td>89.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>C46.ZA</td>
<td>Uncultured bacterium, river sediments (EU262361)</td>
<td>98 Streptomyces yousoufiensis (EN421338)</td>
<td>89.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>C47.ZA</td>
<td>Uncultured bacterium, hot spring, Thailand (AY555773)</td>
<td>96 Acidimicrobium ferrooxidans (CP0001631)</td>
<td>89.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>C50.ZA</td>
<td>Uncultured bacterium, low-level-radioactive-waste site (GQ262855)</td>
<td>97 Frankia alni (CT573213)</td>
<td>88.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>C51.ZA</td>
<td>Uncultured bacterium, environmental sample (EU669645)</td>
<td>98 Acidothermus cellulolyticus (CP0000481)</td>
<td>88.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) Actinobacterial OTUs are defined as < 97% difference. \(^b\) Identification based on the 16S rRNA gene sequence using EzTaxon server (http://www.eztaxon.org)
Table 3. Estimated diversity indices for actinobacterial clone libraries.

<table>
<thead>
<tr>
<th>Location</th>
<th>ACE</th>
<th>Shannon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zambia</td>
<td>9.0</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>(9.0, 9.0)</td>
<td>(1.5, 1.9)</td>
</tr>
<tr>
<td>China</td>
<td>11.4</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>(10.2, 21.0)</td>
<td>(1.6, 2.0)</td>
</tr>
<tr>
<td>New Zealand</td>
<td>0</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.6, 0.7)</td>
</tr>
<tr>
<td>Kenya</td>
<td>10.0</td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td>(10.0, 10.0)</td>
<td>(1.9, 2.2)</td>
</tr>
</tbody>
</table>

Values in parentheses represent the 95% confidence intervals.