Extremophiles (2012) 16:669–679
DOI 10.1007/s00792-012-0465-9

ORIGINAL PAPER

Biogeography of bacterial communities in hot springs: a focus on the actinobacteria

Angel Valverde · Marla Tuffin · Don A. Cowan

Received: 19 April 2012 / Accepted: 25 May 2012 / Published online: 19 June 2012

Abstract Actinobacteria are ubiquitous in soil, freshwater and marine ecosystems. Although various studies have focused on the microbial ecology of this phylum, data are 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 and 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.

Communicated by A. Oren.

Electronic supplementary material The online version of this article (doi:10.1007/s00792-012-0465-9) contains supplementary material, which is available to authorized users.

A. Valverde · M. Tuffin · D. A. Cowan
Institute for Microbial Biotechnology and Metagenomics, University of the Western Cape, Bellville, Cape Town 7535, South Africa

A. Valverde · D. A. Cowan (✉)
University of Pretoria, Pretoria 0002, South Africa
e-mail: don.cowan@up.ac.za

Keywords Actinobacteria · Bacteria · Biogeography · Endemism · Diversity · Hot springs

Introduction

Actinobacteria are ubiquitous highly diverse microorganisms (Allgaier and Grossart 2006; Holmfeldt et al. 2009) 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 (Locci 1994; Trujillo and Goodfellow 2003). Other actinobacterial species form beneficial associations with plants (Benson and Silvester 1993) and insects (Kaltenpoth 2009). Furthermore, actinobacteria are well-known producers of a vast array of secondary metabolites (Watve et al. 2001), 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 (Papke et al. 2003). Hot spring communities are also considered as analogs of early earth and even extraterrestrial conditions (Chapelle et al. 2002). Moreover, thermophiles and their products are of considerable biotechnological interest, having found effective applications in fields as diverse as PCR and biohydrometallurgy (Gericke et al. 2009).

Molecular approaches have revolutionized microbial ecology and revealed that bacteria, like animals and plants,

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

Sample	pH	Temperature (°C)	Cl^- (mg/kg)	NO_3^- (mg/kg)	PO_4^{3-} (mg/kg)	SO_4^{2-} (mg/kg)	Number of TRFs	
							Bacteria	Actinobacteria
ZA-1	5.0	68.3	12	10	31	340	40	10
ZA-2	5.0	59.3	40	7	25	490	32	13
ZA-3	6.0	72.7	50	30	2	700	30	9
CH-1	8.0	66.0	17	6	–	270	34	9
CH-2	7.5	58.8	357.5	17	–	326	34	9
CH-3	8.5	62.5	15	6.5	–	245	32	8
NZ-1	6.5	86.5	1,425	–	140	135	31	9
NZ-2	6.5	86.5	557	–	300	267	31	8
NZ-3	6.5	86.5	3,870	13	50	138	34	8
KE-1	9.5	39.4	50,875	–	–	–	23	16
KE-2	9.5	44.5	43,089	–	–	–	22	16
KE-3	10.0	51.9	27,821	–	–	–	22	17

ZA Zambia, *CH* China, *NZ* New Zealand, *KE* Kenya

exhibit biogeographic patterns shaped by distance and historical and contemporary environmental conditions (reviewed in Martiny et al. 2006).

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 (Huang et al. 2011; Miller et al. 2009; Stott et al. 2008; Whitaker et al. 2003). Such studies have typically focused on the broader bacterial community and very little is known on the biogeography of actinobacteria in hot springs. For instance, Walker et al. (2005) 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. (2009) 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. (2009) has specifically addressed the diversity of this group of microorganisms in hot springs. However, the actinobacterial distribution in these environments has not been related to environmental variables other than temperature. Clearly, more research is needed 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 spring (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 sulfates 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 km wide by 150 km 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 520C 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. (2003) 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 h with agitation (225 rpm). Subsequently, 120 µl of 20 % SDS was added, and the samples were incubated at 65 °C for 1 h. The samples were then centrifuged at 5000×*g* 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 (Berthelet et al. 1996).

Terminal-restriction fragment length polymorphism (T-RFLP) analysis

Bacterial 16S rRNA gene amplification was performed using primer pair 341F (5′-CCTACGGGAGGCAGCAI-3′) (Ishii and Fukui 2001), and 908R (5′-CCGTCAATTCMT TTGAGTTI-3′) (Lane et al. 1985). For actinobacterial-specific amplifications, primers 226–243F (5′-GGATGAG CCCGCGGCCTA-3′) (Heuer et al. 1997) and (A3R) 1414–1430R (5-CCAGCCCCACCTTCGAC-3) (Monciardini et al. 2002) were used. Bacterial primers were modified by the addition of inosine at the 3′ end in an attempt to broaden their target scope (Ben-Dov et al. 2006). In both primer pairs, the forward primer was labeled with 6′ carboxyfluorescein (6-FAM).

Each PCR mixture (50 µl) contained 2 U 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 *Taq*I, *Alu*I and *Hha*I 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. (2006). 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 (Clarke and Warwick 2001). 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 (Clarke and Gorley 2006). 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 (Ramette 2007). The influence of environmental factors on the community structure of the samples was assessed using the BEST analysis in PRIMER 6. The BEST analysis selects the best subset of environmental variables, so that the Euclidean distances of scaled environmental variables have the maximum correlation with community dissimilarities (i.e., the highest Spearman's rho). The significance of the correlations between two distance matrices was calculated using the Mantel test (Legendre and Legendre 1998) 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

Clone libraries were constructed after pooling equal amounts of amplicons from the individual samples for each location obtained with 226–243F and 1414–1430R 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 *Alu*I and *Rsa*I) was used to de-replicate clones. Restriction patterns were visualized on 2 % agarose gels and analyzed using Gelcompare II (Applied Maths, Keistraat, Belgium). Plasmid DNA, from a representative of each unique restriction pattern, was extracted with QIAprep Spin Miniprep kit (Qiagen GmbH, Hilden, Germany) and sequenced using the vector primer M13F with an ABI 3130 DNA Sequencer (Applied Biosystems).

Putative chimeric sequences were filtered using Bellerophon (Hubert et al. 2009). Sequences of >97 % identity were grouped into OTUs using MOTHUR (Schloss et al. 2009). Taxonomic assignments of representative OTUs were obtained using the Classifier tool (Wang et al. 2007) (confidence threshold 80 %) from the Ribosomal Database Project II (http://rdp.cme.msu.edu/) (Cole et al. 2009). Sequences were aligned using the SINA aligner from the SILVA project (http://www.arb-silva.de) (Pruesse et al. 2007) and imported into ARB (Ludwig et al. 2004) where alignments were manually curated. The phylogenetic tree was constructed by the neighbor-joining method (Saitou and Nei 1987) after the Jukes–Cantor model for nucleotide substitution was applied. The robustness of the tree topology was evaluated by bootstrap analysis (Felsenstein 1985) 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 % cutoff (species level) (Schloss and Handelsman 2004).

Intercommunity OTU composition was evaluated with UniFrac analysis (http://bmf2.colorado.edu/unifrac/) (Lozupone et al. 2006). Jackknife analysis was used to assess the confidence of tree nodes. Clone libraries were compared using ∫-Libshuff analysis (Schloss et al. 2004).

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 (KE).

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 (ESM S1). 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 (Fig. 1). Both bacteria and actinobacteria showed a clear separation of samples according to their geographic origin. 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 (ρ) 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.

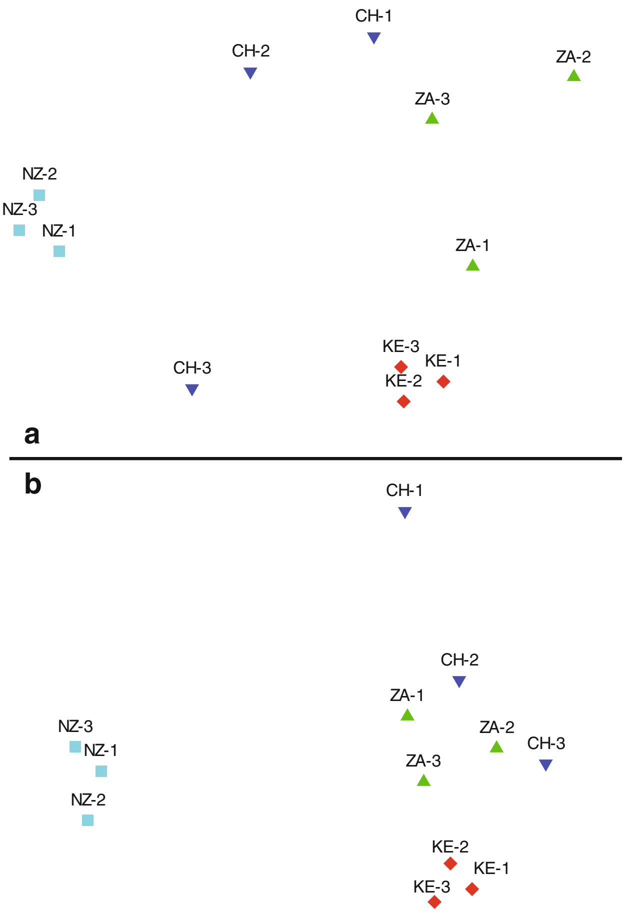


Fig. 1 Non-metric multidimensional scaling based on T-RFLP profiles. **a** Actinobacteria (stress = 0.13); **b** bacteria (stress = 0.09). *ZA* Zambia, *CH* China, *NZ* New Zealand, *KE* Kenya

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 $P \leq 0.05$).

Actinobacterial 16S rRNA gene library analysis

A total of 400 actinobacterial clones (100 per library) were subjected to RFLP analysis, identifying 51 unique banding patterns (data not shown). A total of 28 different OTUs (97 % similarity cutoff) 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; Fig. 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 (Fig. 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 indicate 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 ∫-Libshuff analysis of the 16S rRNA gene clone libraries indicated that the communities differed significantly between all pairwise combinations ($P \leq 0.05$, 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 (ESM S2).

Discussion

Molecular fingerprinting methods, such as T-RFLP, allow descriptions of broad-scale changes in the taxonomic composition of bacterial communities (Liu et al. 1997) although the limitations and disadvantages of the technique are well documented (e.g., Zhang et al. 2008).

Our results are in agreement with Zhang et al. (2008) 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 (Lage et al. 2010; Zhang et al. 2008).

In the last decade, information gathered from molecular surveys suggests that microbial biogeographic patterns are shaped by history and environmental factors (Martiny et al. 2006). 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 (Fierer and Jackson 2006). Miller et al. (2009), analyzing two hot spring bacterial communities along the temperature gradients in YNP, showed that diversity

Table 2 16S rRNA clones identified in hot springs sediments

OTU[a]	Representative sequence	Accession no.	Closest sequence match with BLASTN, origin (accession number)	Similarity (%)	Closest type strain[b]	Similarity (%)
1	C1_KE	JN806355	*Couchiplanes caeruleus*, soil (NR_026295)	99	*Couchiplanes caeruleus* (X93202)	99.3
2	C2_KE	JN806356	*Glycomyces harbinensis*, soil (AJ293747)	98	*Glycomyces harbinensis* (D85483)	98.1
3	C3_KE	JN806357	Uncultured bacterium, soil (FJ478553)	96	*Streptomyces albiaxialis* (AY999901)	87.5
4	C4_KE	JN806358	Uncultured bacterium, lead–zinc mine (EF612367)	98	*Frankia alni* (CT573213)	88.3
5	C6_KE	JN806359	*Mycobacterium* sp., tap water (GU084181)	99	*Mycobacterium parascrofulaceum* (ADNV01000350)	99.1
6	C7_KE	JN806360	Uncultured bacterium, Hanford 300 area (HM186589)	95	*Thermosipho altlanticus* (AJ577471)	51.0
7	C8_KE	JN806361	*Aeromicrobium* sp., soil (JF824798)	99	*Aeromicrobium panaciterrae* (AB245387)	97.6
8	C10_KE	JN806362	Uncultured bacterium, air dust (EF683032)	98	*Nocardioides iriomotensis* (AB544079)	96.9
9	C11_KE	JN806363	Uncultured *Modestobacter*, Taklamatan desert (JF411374)	96	*Modestobacter versicolor* (AJ871304)	96.1
10	C12_KE	JN806364	Uncultured bacterium, soil contaminated with anthracene (HM438001)	97	*Aciditerrimonas ferrireducens* (AB517669)	88.5
11	C14_NZ	JN806365	Uncultured bacterium, saline-alkaline soil (JN037894)	96	*Streptomyces glaucosporus* (AB184664)	89.1
12	C15_NZ	JN806366	Uncultured bacterium, geothermal system, Yellowstone National Park (DQ324877)	99	*Nitrosococcus oceanis* (AY033493)	87.3
13	C21_CH	JN806367	Uncultured bacterium, Hanford 300 area (HM186978)	95	*Laceyella putida* (AF138736)	49.3
14	C23_CH	JN806368	Uncultured actinobacterium, sludge (CU924270)	97	*Kribbella yunnanensi* (CP000127)	49.3
15	C28_CH	JN806369	Uncultured bacterium, radioactive-waste site (GQ263231)	95	*Thermoleophilum minutum* (AJ458464)	88.3
16	C25_CH	JN806370	Uncultured bacterium, prairie soil (EU132603)	96	*Streptomyces panacagri* (AB245388)	89.8
17	C29_CH	JN806371	Uncultured bacterium, soil (EU132572)	90	*Frankia alni* (CT573213)	50.1
18	C30_CH	JN806372	Uncultured bacterium, hot spring, Thailand (AY555773)	93	*Thermoleophilum minutum* (AJ458464)	48.2
19	C32_CH	JN806373	Uncultured bacterium, prairie soil (FJ479127)	96	*Thermotoga subterranea* (U22664)	48.9
20	C33_CH	JN806382	Uncultured bacterium, soil (EF020292)	96	*Actinomadura cremea* (AF134067)	85.3
21	C37_CH	JN806374	Uncultured bacterium, rizhosphere soil (EU786138)	98	*Acidothermus cellulolyticus* (CP000481)	89.5
22	C39_ZA	JN806375	Uncultured bacterium, soil contaminated with anthracene (HM438001)	99	*Streptomyces capillispiralis* (AB184577)	88.4
23	C40_ZA	JN806376	Uncultured bacterium, soil (EF516773)	96	*Luedemannella flava* (AB236959)	88.2
24	C45_ZA	JN806377	Uncultured bacterium, soil (FJ478871)	96	*Streptomyces caniferus* (AB184640)	89.1
25	C46_ZA	JN806378	Uncultured bacterium, river sediments (EU262361)	98	*Streptomyces youssoufienis* (FN421338)	89.8
26	C47_ZA	JN806379	Uncultured bacterium, hot spring, Thailand (AY555773)	96	*Acidimicrobium ferrooxidans* (CP001631)	89.1
27	C50_ZA	JN806380	Uncultured bacterium, low-level-radioactive-waste site (GQ262855)	97	*Frankia alni* (CT573213)	88.4
28	C51_ZA	JN806381	Uncultured bacterium, environmental sample (EU669645)	98	*Acidothermus cellulolyticus* (CP000481)	88.0

[a] Actinobacterial OTUs are defined at 97 % similarity values

[b] Identification based on the 16S rRNA gene sequence using EzTaxon server (http://www.eztaxon.org)

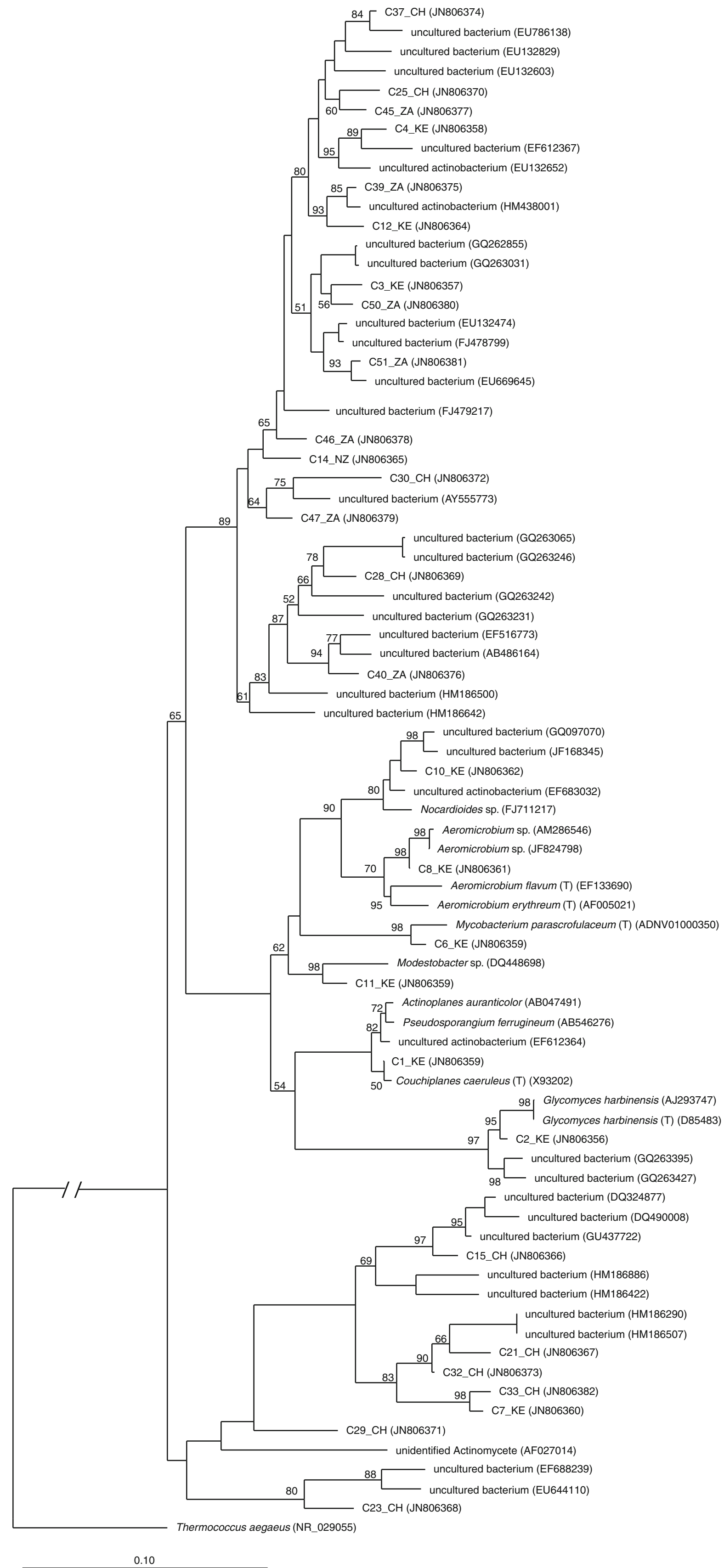


Fig. 2 Neighbor-joining tree of representative 16S rRNA sequences described in this study. Only bootstrap values ≥50 % are shown. *Bar* 0.1 substitutions per nucleotide position

Table 3 Estimated diversity indices for actinobacterial clone libraries

Location	ACE	Shannon
Zambia	9.0 (9.0, 9.0)	1.7 (1.5, 1.9)
China	11.4 (10.2, 21.0)	1.8 (1.6, 2.0)
New Zealand	0	0.7 (0.6, 0.7)
Kenya	10.0 (10.0, 10.0)	2.1 (1.9, 2.2)

Values in parentheses represent the 95 % confidence intervals

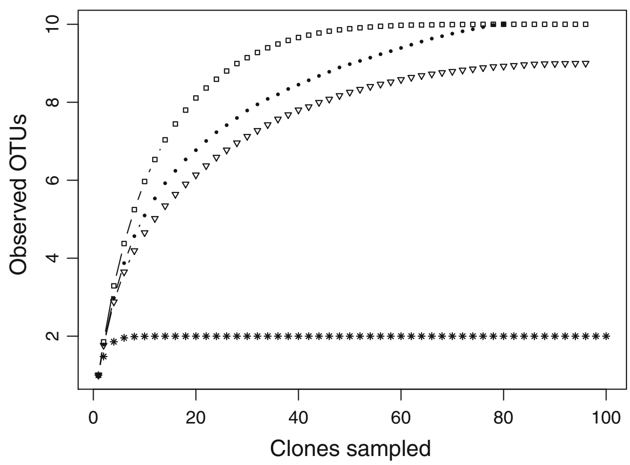


Fig. 3 Rarefaction curves indicating the observed number of OTUs at a genetic distance of 3 % in hot springs sediments. Kenya (*squares*), Zambia (*triangles*), China (*circles*), New Zealand (*asterisks*)

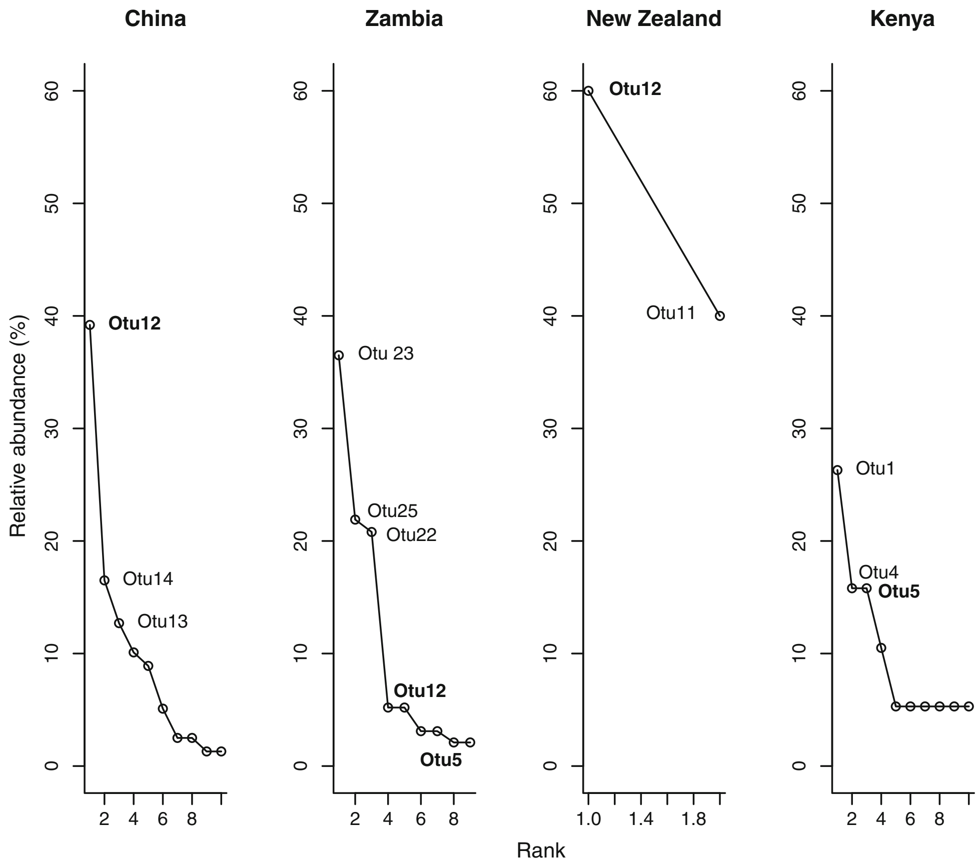
decreased with increasing temperature. In addition, they found that community similarity decayed exponentially with increasing differences in temperature between samples but not with distance. In contrast, Papke et al. (2003) demonstrated that distance effects dominated cyanobacterial diversity in hot 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. This is not surprising, as different groups of organisms may respond differently to the same local environment (Drakare and Liess 2010). Using T-RFLP and clone library analyses, we found that actinobacterial richness and diversity decreased with increasing temperature. In contrast to Miller et al. (2009), we did not observe a decrease in the number of bacterial OTUs at the site with the highest temperature, and as it has been previously reported (Purcell et al. 2007), 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 (Martiny et al. 2011). Nevertheless, as can be inferred from Table 1 and rarefaction curves (Fig. 3), and has been shown in other studies (e.g. Bent and Forney 2008), some discrepancy exists in the number of actinobacterial OTUs observed using both methodologies. Clearly, further research is required 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 (Papke et al. 2003; Purcell et al. 2007; Whitaker et al. 2003). 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 (ESM S2). Taken together, these findings, and those from Song et al. (2009), indicate that the differences in actinobacterial community composition in hot springs are a function of both environmental heterogeneity and geographic distance (Martiny et al. 2006, 2011).

Phylogenetic analyses revealed a high diversity of actinobacterial OTUs in hot springs samples, the majority of sequences belonging to unclassified taxa. Only 4 of 28 OTUs showed high homology (>97 %) to known actinobacterial species: *Couchiplanes caeruleus* (OTU 1), *Glycomyces harbinensis* (OTU 2) *Aeromicrobium panaciterrae* (OTU 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 (Santos et al. 2007). 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 (Miller et al. 2009; Song et al. 2009). However, sequences from "extreme" environments are poorly represented in these

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



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 (Fig. 2). In agreement with these findings, Song et al. (2009) 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 Armatimonadetes). 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 (ESM S1, Fig. 4). Moreover, the phylotype abundance distribution showed the dominance of a few OTUs in each library (Fig. 4), whereas most of the OTUs were relatively rare, resembling the classic "long tail" phenomenon (Fuhrman 2009). 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) (Fig. 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 (Papke et al. 2003; Whitaker et al. 2003) 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 are supported by the recent work by Nemergut et al. (2011), 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.

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

References

Abdo Z, Schuette UME, Bent SJ, Williams CJ, Forney LJ, Joyce P (2006) Statistical methods for characterizing diversity of microbial communities by analysis of terminal restriction fragment length polymorphisms of 16S rRNA genes. Environ Microbiol 8:929–938

Allgaier M, Grossart HP (2006) Diversity and seasonal dynamics of actinobacteria populations in four lakes in northeastern Germany. Appl Environ Microbiol 72:3489–3497

Ben-Dov E, Shapiro OH, Siboni N, Kushmaro A (2006) Advantage of using inosine at the 3′ termini of 16S rRNA gene universal primers for the study of microbial diversity. Appl Environ Microbiol 72:6902–6906

Benson DR, Silvester WB (1993) Biology of Frankia strains, actinomycete symbionts of actinorhizal plants. Microbiol Rev 57:293–319

Bent SJ, Forney LJ (2008) The tragedy of the uncommon: understanding limitations in the analysis of microbial diversity. ISME J 2:689–695

Berthelet M, Whyte LG, Greer CW (1996) Rapid, direct extraction of DNA from soils for PCR analysis using polyvinylpolypyrrolidone spin columns. FEMS Microbiol Lett 138:17–22

Chapelle FH, O'Neill K, Bradley PM, Methe BA, Ciufo SA, Knobel LL, Lovley DR (2002) A hydrogen-based subsurface microbial community dominated by methanogens. Nature 415:312–315

Clarke KR, Gorley RN (2006) PRIMER v6: user manual. Plymouth Marine Laboratory, Plymouth

Clarke KR, Warwick RM (2001) Changes in marine communities: an approach to statistical analysis and interpretation, 2nd edn. Primer-E Ltd.: Plymouth, Marine Laboratory, Plymouth

Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ, Kulam-Syed-Mohideen AS, McGarrell DM, Marsh T, Garrity GM, Tiedje JM (2009) The ribosomal database project: improved alignments and new tools for rRNA analysis. Nucleic Acids Res 37:D141–D145

Drakare S, Liess A (2010) Local factors control the community composition of cyanobacteria in lakes while heterotrophic bacteria follow a neutral model. Freshwater Biol 55:2447–2457

Ellis RJ, Morgan P, Weightman AJ, Fry JC (2003) Cultivation-dependent and -independent approaches for determining bacterial diversity in heavy-metal-contaminated soil. Appl Environ Microbiol 69:3223–3230

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Fierer N, Jackson RB (2006) The diversity and biogeography of soil bacterial communities. Proc Natl Acad Sci USA 103:626–631

Fuhrman JA (2009) Microbial community structure and its functional implications. Nature 459:193–199

Gericke M, Neale JW, van Staden PJ (2009) A Mintek perspective of the past 25 years in minerals bioleaching. J S Afri Inst Min Metall J 109:567–585

Heuer H, Krsek M, Baker P, Smalla K, Wellington EMH (1997) Analysis of actinomycete communities by specific amplification of genes encoding 16S rRNA and gel-electrophoretic separation in denaturing gradients. Appl Environ Microbiol 63:3233–3241

Holmfeldt K, Dziallas C, Titelman J, Pohlmann K, Grossart H-P, Riemann L (2009) Diversity and abundance of freshwater actinobacteria along environmental gradients in the brackish northern Baltic Sea. Environ Microbiol 11:2042–2054

Huang Q, Dong CZ, Dong RM, Jiang H, Wang S, Wang G, Fang B, Ding X, Niu L, Li X, Zhang C, Dong H (2011) Archaeal and bacterial diversity in hot springs on the Tibetan Plateau, China. Extremophiles 15:549–563

Hubert C, Loy A, Nickel M, Arnosti C, Baranyi C, Brüechert V, Ferdelman T, Finster K, Christensen FM, de Rezende JR, Vandieken V, Jørgensen BB (2009) A constant flux of diverse thermophilic bacteria into the cold arctic seabed. Science 325:1541–1544

Ishii K, Fukui M (2001) Optimization of annealing temperature to reduce bias caused by a primer mismatch in multitemplate PCR. Appl Environ Microbiol 67:3753–3755

Kaltenpoth M (2009) Actinobacteria as mutualists: general healthcare for insects? Trends Microbiol 17:529–535

Lage MD, Reed HE, Weihe C, Crain CM, Martiny JBH (2010) Nitrogen and phosphorus enrichment alter the composition of ammonia-oxidizing bacteria in salt marsh sediments. ISME J 4:933–944

Lane DJ, Pace B, Olsen GJ, Stahl DA, Sogin ML, Pace NR (1985) Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. Proc Natl Acad Sci USA 82:6955–6959

Legendre P, Legendre L (1998) Numerical ecology. Elsevier Science BV, Amsterdam

Liu WT, Marsh TL, Cheng H, Forney LJ (1997) Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. Appl Environ Microbiol 63:4516–4522

Locci R (1994) Actinomycetes as plant pathogens. Eur J Plant Pathol 100:179–200

Lozupone C, Hamady M, Knight R (2006) UniFrac: an online tool for comparing microbial community diversity in a phylogenetic context. BMC Bioinform 7:371

Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar, Buchner A, Lai T, Steppi S, Jobb G, Forster W, Brettske I, Gerber S, Ginhart AW, Gross O, Grumann S, Hermann S, Jost R, Konig A, Liss T, Lussmann R, May M, Nonhoff B, Reichel B, Strehlow R, Stamatakis A, Stuckmann N, Vilbig A, Lenke M, Ludwig T, Bode A, Schleifer KH (2004) ARB: a software environment for sequence data. Nucleic Acids Res 32:1363–1371

Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL, Horner-Devine MC, Kane M, Krumins JA, Kuske CR, Morin PJ, Naeem S, Øvreås L, Reysenbach AL, Smith VH, Staley JT (2006) Microbial biogeography: putting microorganisms on the map. Nat Rev Microbiol 4:102–112

Martiny JBH, Eisen JA, Penn K, Allison SD, Horner-Devine MC (2011) Drivers of bacterial beta-diversity depend on spatial scale. Proc Natl Acad Sci USA 108:7850–7854

Miller SR, Strong AL, Jones KL, Ungerer MC (2009) Bar-coded pyrosequencing reveals shared bacterial community properties along the temperature gradients of two alkaline hot springs in Yellowstone National Park. Appl Environ Microbiol 75:4565–4572

Monciardini P, Sosio M, Cavaletti L, Chiocchini C, Donadio S (2002) New PCR primers for the selective amplification of 16S rDNA from different groups of actinomycetes. FEMS Microbiol Ecol 42:419–429

Nemergut DR, Costello EK, Hamady M, Lozupone C, Jiang L, Schmidt SK, Fierer N, Townsend AR, Cleveland CC, Stanish L, Knight R (2011) Global patterns in the biogeography of bacterial taxa. Environ Microbiol 13:135–144

Papke RT, Ramsing NB, Bateson MM, Ward DM (2003) Geographical isolation in hot spring cyanobacteria. Environ Microbiol 5:650–659

Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig WG, Peplies J, Glöckner FO (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 35:7188–7196

Purcell D, Sompong U, Yim LC, Barraclough TG, Peerapornpisal Y, Pointing SB (2007) The effects of temperature, pH and sulphide on the community structure of hyperthermophilic streamers in hot springs of northern Thailand. FEMS Microbiol Ecol 60:456–466

Ramette A (2007) Multivariate analyses in microbial ecology. FEMS Microbiol Ecol 62:142–160

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Santos R, Fernandes J, Fernandes N, Oliveira F, Cadete M (2007) *Mycobacterium parascrofulaceum* in acidic hot springs in Yellowstone National Park. Appl Environ Microbiol 73:5071–5073

Schloss PD, Handelsman J (2004) Status of the microbial census. Microbiol Mol Biol Rev 68:686–691

Schloss PD, Larget BR, Handelsman J (2004) Integration of microbial ecology and statistics: a test to compare gene libraries. Appl Environ Microbiol 70:5485–5492

Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol 75:7537–7541

Song Z, Zhi X, Li W, Jiang H, Zhang C, Dong H (2009) Actinobacterial diversity in hot springs in Tengchong (China), Kamchatka (Russia), and Nevada (USA). Geomicrobiol J 26:256–263

Stott MB, Crowe MA, Mountain BW, Smirnova AV, Hou S, Alam M, Dunfield PF (2008) Isolation of novel bacteria, including a candidate division, from geothermal soils in New Zealand. Environ Microbiol 10:2030–2041

Trujillo ME, Goodfellow M (2003) Numerical phenetic classification of clinically significant aerobic sporoactinomycetes and related organisms. Antonie Van Leeuwenhoek 84:39–68

Walker JJ, Spear JR, Pace NR (2005) Geobiology of a microbial endolithic community in the Yellowstone geothermal environment. Nature 434:1011–1014

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 73:5261–5267

Watve MG, Tickoo R, Jog MM, Bhole BD (2001) How many antibiotics are produced by the genus *Streptomyces*? Arch Microbiol 176:386–390

Whitaker RJ, Grogan DW, Taylor JW (2003) Geographic barriers isolate endemic populations of hyperthermophilic archaea. Science 301:976–978

Zhang R, Thiyagarajan V, Qian P-Y (2008) Evaluation of terminal-restriction fragment length polymorphism analysis in contrasting marine environments. FEMS Microbiol Ecol 65:169–178