The ISME Journal (2010) 4, 346–356

www.nature.com/ismej

ORIGINAL ARTICLE

Diversity and abundance of *Korarchaeota* in terrestrial hot springs of Iceland and Kamchatka

Laila Johanne Reigstad[1], Steffen Leth Jorgensen[1] and Christa Schleper[1,2]
[1]*Centre for Geobiology and Department of Biology, University of Bergen, Allégaten, Bergen, Norway and* [2]*Department of Genetics in Ecology, University of Vienna, Althanstrasse, Vienna, Austria*

Korarchaeota **constitute a recently proposed and little characterized kingdom of Archaea that might have diverged before the lineages of Crenarchaeota and Euryarchaeota split. To assess the diversity, distribution and abundance of** ***Korarchaeota*****, we analysed 19 terrestrial hot springs in Hveragerdi and Krysuvik, Iceland, and in Kamchatka, Russia. The springs were 70–97 °C with pH 2.5–6.5. Out of 19 springs, 12 tested positive for** ***Korarchaeota*** **with specific primers. A** ***Korarchaeota*** **16S rDNA library was made from each of these. From the 301 clones sequenced, 87 unique sequences were obtained from Iceland and 33 from Kamchatka. The similarity between Kamchatkan and Icelandic 16S rDNA sequences and that of** ***Candidatus Korarchaeum cryptofilum*** **was ⩽93.5%. Phylogenetic analyses revealed a clear separation between sequences retrieved from terrestrial and marine habitats. Within the terrestrial sequences, four clusters could be recognized showing a geographic distribution with surprisingly low diversity. Furthermore, the abundance of** ***Korarchaeota*** **16S rDNA in the 12 environmental samples was analysed using quantitative PCR (qPCR), showing that** ***Korarchaeota*** **represent only a minor fraction of the microbial community in hot springs; however, in some cases they constitute up to 7% of all Archaea. Taxonomic profiling of an Icelandic** ***Korarchaeota*****-positive habitat revealed an** ***Aquificales*****-dominated community. In fact,** ***Aquificales*** **were dominating or present in high numbers in all 12 positive sites. Chemical analyses of three** ***Korarchaeota*****-positive hot springs showed their occurrence in variable water chemistry. Our data provide new information on** ***Korarchaeota*** **habitats and shed light on their abundance, diversity, distribution and coexisting organisms.**
The ISME Journal (2010) **4,** 346–356; doi:10.1038/ismej.2009.126; published online 3 December 2009
Subject Category: microbial population and community ecology
Keywords: *Aquificales*; hot spring; *Korarchaeota*; qPCR; Iceland; Kamchatka

Introduction

Since their recognition as a separate domain of life, Archaea have been interesting study objects in the context of early evolution and early life forms on Earth. Not only their phylogenetic position and their specialization for extreme living conditions, but also their unique and often eukaryotic-like cellular features have raised interest (Garrett and Klenk, 2006; Cavicchioli, 2007). Extensive 16S rRNA gene surveys and phylogenetic analyses have recently led to the discovery of several deep-branching lineages within the Archaea that might be particularly helpful in dissecting features of the last common ancestor that preceded all prokaryotic and eukaryotic life form (Barns *et al.*, 1994, 1996; Takai and Horikoshi, 1999; Huber *et al.*, 2002). These deep-branching lineages include the *Korarchaeota* kingdom (Barns *et al.*, 1994, 1996), the deep-sea Archaeal group, the marine hydrothermal vent group (Takai *et al.*, 2001) (K Takai, personal communication), and perhaps (although debated) the *Nanoarchaeota* kingdom (Huber *et al.*, 2002), or the newly proposed *Thaumarchaeota* kingdom (Brochier-Armanet *et al.*, 2008). Among the deep-branching lineages, in particular, *Korarchaeota* is considered a kingdom that might have diverged very early in the evolution of Archaea (Elkins *et al.*, 2008), even before the *Euryarchaeota* and *Crenarchaeota* split (Barns *et al.*, 1994, 1996). Until recently, the analyses of *Korarchaeota* were restricted to 16S rRNA-based molecular approaches and a few enrichment cultures (Barns *et al.*, 1994, 1996; Burggraf *et al.*, 1997; Stetter, 2006), as no organisms so far have been obtained in pure laboratory culture. However, most recently Elkins *et al.* (2008) presented the first korarchaeal genome of *Candidatus Korarchaeum cryptofilum*, being physically isolated from an enrichment culture inoculated with sediments from Obsidian Pool, Yellowstone. During continuous anaerobic culturing at 85 °C and pH 6.5, the organism appeared as ultrathin long filamentous cells. The information retrieved from

Correspondence: LJ Reigstad, Centre for Geobiology and Department of Biology, University of Bergen, Allégaten 41, Bergen, N-5007, Norway.
E-mail: laila.reigstad@bio.uib.no
Received 25 June 2009; revised and accepted 27 October 2009; published online 3 December 2009

the genome indicates that the organism is a proficient peptide degrader and obtains its carbon and energy through fermentation. It is apparently unable to synthesize essential co-factors, vitamins and purines and is therefore suggested to be either a symbiotic organism or a very efficient scavenger (Nealson, 2008).

Since the discovery of *Korarchaeota* more than a decade ago (Barns *et al.*, 1996; Burggraf *et al.*, 1997) several environmental 16S rRNA gene surveys have been conducted, identifying *Korarchaeota* in terrestrial hot springs (Barns *et al.*, 1994, 1996; Takai and Sako, 1999; Reysenbach *et al.*, 2000; Skirnisdottir *et al.*, 2000; Hjorleifsdottir *et al.*, 2001; Spear *et al.*, 2005; Auchtung *et al.*, 2006; Elkins *et al.*, 2008), in shallow marine hydrothermal vents (Marteinsson *et al.*, 2001a, b), in hydrothermal wells (Rogers and Amend, 2005), in deep-sea hydrothermal vent fluids (Takai and Sako, 1999; Nercessian *et al.*, 2003), sediments (Teske *et al.*, 2002), chimneys (Takai *et al.*, 2004; Auchtung *et al.*, 2006) and on the sea floor of abyssal hills (Ehrhardt *et al.*, 2007a). *Korarchaeota* 16S rDNA signatures are found in thermal habitats of 55–128 °C and mostly at (sub-) neutral pH. By relating the presence and diversity of *Korarchaeota* with geographic locations, five different phylogenetic groups have been suggested with a possible correlation between diversity and geography (Auchtung *et al.*, 2006).

Although the description of *Candidatus Korarchaeum cryptofilum* (Elkins *et al.*, 2008) brings us much closer to the organisms of the *Korarchaeota* kingdom, especially concerning physiology, there are still unanswered questions regarding their ecological role and abundance. Further, it is an open question whether the information retrieved from this particular organism holds true for the whole kingdom. To address some of these questions, we have surveyed the occurrence and abundance of *Korarchaeota* in the frame of a broader meta-genomic and phylogenetic study of terrestrial hot springs. Our 16S rRNA-based studies expand the known korarchaeal diversity and give insights into their abundance. In addition, we provide data on korarchaeal habitats that might assist in future meta-genomic analyses and cultivation attempts.

Materials and methods

Sampling

Samples were collected from terrestrial hot springs ($\geqslant$70 °C) in Kamchatka Peninsula, Russia, in August 2005, and from Hveragerdi and Krysuvik, Iceland in April 2006 and September 2007.

On the Kamchatka Peninsula, samples were taken from seven hot springs: three from the Uzon Caldera and four from Mt Mutnovsky hydrothermal field. At Hveragerdi, 10 hot spring samples were collected from Hverakjalki (one site), Reykjadalur (seven sites) and Grensdalur (two sites). At Krysuvik, two hot springs were sampled at Seltun. Different kinds of hot spring material were harvested: filament samples refer to the fibrous fingerlong structures attached to the edges of some hot springs, mud samples were harvested in the middle of the hot springs below the surface and the more solid greyish microbial mats were harvested from the edge of two hot springs, but still well below water surface. Descriptions of the 19 different springs investigated for the presence of *Korarchaeota* 16S rRNA genes are given in Table 1.

DNA extraction, PCR amplification and sequencing

Kamchatkan samples were kept in the dark at 6 °C for 3–5 days until DNA extraction was performed, as freezing was not possible because of the remoteness of sampling sites. DNA extraction from Icelandic samples was either carried out within 2 h after sampling or samples were frozen at −20 °C for later processing. DNA was extracted from Kamchatkan and Icelandic samples using FastPrep DNA Extraction Kit for soil (Bio 101, Q-BIOgene, Illkirch, France). The protocol of the manufacturer was followed, starting with 0.5 g sample material using a FastPrep instrument setting of 5.5 for 30 s. *Korarchaeota* 16S rRNA genes were amplified by PCR using primers Kor236F and Kor1236R (Auchtung *et al.*, 2006). *Aquificales* 16S rRNA genes were PCR amplified using primers aqfx540modF (5′-GGTCCCGARCGTTGCGCGA-3′) being our modified version of aqfx540R (Reysenbach *et al.*, 2005), and primer Aqui1197R, originally used as an *Aquificales*-specific probe (Rusch and Amend, 2004). Specificity of these *Aquificales* primers was confirmed by cloning and sequencing of amplicons (not shown). The *Korarchaeota* PCR was performed at 95 °C for 5 min and 30 cycles at 95 °C for 45 s, 59 °C for 45 s and 72 °C for 70 s, followed by 72 °C for 10 min and cooling to 4 °C. The *Aquificales* PCR was conducted at 95 °C for 5 min and 25 cycles at 95 °C for 30 s, 62 °C for 30 s and 72 °C for 30 s, followed by 72 °C for 10 min and cooling to 4 °C. Both PCR reactions (20 μl) contained 0.17 mM each deoxynucleotide triphosphate, 1.75 mM magnesium chloride, 0.7 mM each primer, 1 M betaine (Sigma, St Louis, MO, USA), 1 × PCR buffer, 0.5 U *Taq* polymerase (Fermentas, Hanover, MD, USA) and approximately 10 ng template DNA.

A clone library from each of the 12 *Korarchaeota*-positive hot springs was constructed by cloning *Korarchaeota* 16S rDNA PCR amplicon into pCR4-TOPO and transforming into TOP10 *Escherichia coli* cells (*E. coli*) (Invitrogen, Carlsbad, CA, USA). Colonies were screened by PCR using vector primers M13F and M13R to identify clones with appropriately sized inserts. DNA inserts from a total of 301 randomly chosen clones (Table 1) were cleaned by GeneElute PCR Clean-Up Kit (Sigma) and sequenced using BigDye v3.1 chemistry and run on ABI 3730xl sequencer (Applied Biosystems, Carlsbad, CA, USA).

Table 1 Description of the 19 terrestrial hot springs screened for *Korarchaeota* 16S rRNA genes

Site	*Location*	*Country*	*Sample ID*	*°C*	*pH*	*TOC*[a]*(%)*	*Korarchaeota 16S rDNA*[b]	*Aquificales 16S rDNA*[c]
IceG1	Reykjadalur	Iceland	Filaments	81	5.5	0.20 (0.36)	Yes (33)	Yes
IceG2	Reykjadalur	Iceland	Microbial mat	81	5.5	0.21 (1.10)	Yes (29)	Yes
IceG3	Reykjadalur	Iceland	Mud	85	3.0	0.08 (0.19)	No	Yes
IceG8	Hverakjalki	Iceland	Mud	85	3.5	0.02 (0.06)	No	No
IceG13	Reykjadalur	Iceland	Filaments	84	5.5	0.43 (1.16)	Yes (31)	Yes
IceG13A	Reykjadalur	Iceland	Filaments	97	6.5	NA	Yes (33)	Yes
IceG14	Reykjadalur	Iceland	Filaments	70	5.0	0.46 (1.40)	Yes (33)	Yes
IceG16	Reykjadalur	Iceland	Filaments	82	5.0	0.16 (0.78)	Yes (28)	Yes
IceR6	Grensdalur	Iceland	Mud	83	3.0	0.01 (0.21)	No	No
IceR11	Grensdalur	Iceland	Microbial mat	81	3.0	0.04 (0.21)	No	Yes
Seltun1	Krysuvik	Iceland	Filaments	76	6.5	NA	Yes (31)	Yes
Seltun3	Krysuvik	Iceland	Filaments	74	6.0	NA	Yes (33)	Yes
Kam36	Uzon Caldera	Russia	Filaments	80	6.0	0.16 (1.85)	Yes (33)	Yes
Kam38	Uzon Caldera	Russia	Filaments	82	5.5	0.12 (0.56)	Yes (41)	Yes
Kam40	Uzon Caldera	Russia	Filaments	83	6.0	0.05 (0.34)	Yes (34)	Yes
Kam41	Mt Mutnovsky	Russia	Mud	94	2.5	0.04 (0.13)	No	No
Kam42	Mt Mutnovsky	Russia	Mud	94	4.5	0.01 (0.54)	No	No
Kam43	Mt Mutnovsky	Russia	Mud	85	2.5	0.03 (0.20)	No	No
Kam44	Mt Mutnovsky	Russia	Filaments	84	6.5	0.03 (0.16)	Yes (39)	Yes

Abbreviation: NA, not analysed.
[a]Total organic carbon (TOC) estimated as percentage of organic carbon per gram wet weight sample. Number in brackets is percentage of TOC of dry weight.
[b]Based on *Korarchaeota*-specific 16S rDNA PCR, cloning and sequencing. Numbers in brackets refer to number of sequenced clones.
[c]Based on *Aquificales*-specific 16S rDNA PCR.

Real-time quantitative PCR (qPCR)

Korarchaeota 16S rRNA genes were enumerated using specific primers (Kor236F/Kor1236R). Each reaction (20 μl) contained 1 × QuantiTech Sybr Green PCR master mixture (Qiagen, Venlo, the Netherlands), 0.5 mM of each primer and 1 μl template DNA (approximately 10 ng DNA). The thermal cycling program was 15 min at 95 °C, and then 40 cycles of 94 °C for 30 s, 59 °C for 20 s, 72 °C for 45 s and 84 °C for 2 s. Quantification standard consisted of a dilution series of purified DNA amplicon from a conventional PCR using the Kor236F/Kor1236R as described previously with an IceG13 *Korarchaeota* 16S rDNA-positive clone as template. Copy number of *Korarchaeota* 16S rRNA genes in the standard ranged between 17 and 1.7×10^6 copies μl^{-1}. Genomic DNA from *E. coli* was used as negative control. The R^2 value for standard curve was 0.99 and slope value was −3.28, indicating a PCR amplification efficiency of 98%.

Archaeal 16S rRNA genes were quantified using primers 20F/958R (DeLong, 1992). Each reaction (20 μl) contained 1 × QuantiTech Sybr Green PCR master mixture (Qiagen), 0.5 mM of each primer and 1 μl template DNA (approximately 10 ng DNA). Thermal cycling program was: 15 min at 95 °C, then 40 cycles of 94 °C for 30 s, 62 °C for 30 s, 72 °C for 45 s and 80 °C for 2 s. Quantification standard consisted of a dilution series of a known amount of linearized fosmid 54d9 (Treusch *et al.*, 2005) and a copy number of Archaeal 16S rRNA genes between 10 and 1×10^7 copies μl^{-1} (calculated as described by Leininger *et al.*, 2006). Genomic DNA from *E. coli* was used as negative control. The R^2 value for the standard curve was 0.99 and slope value was −3.71, giving an estimated amplification efficiency of 86%.

Bacterial 16S rRNA genes were quantified using primers 338F/519R and following the protocol described by Einen *et al.* (2007). Quantification standard consisted of a dilution series (between 1×10^2 and 1×10^7 copies μl^{-1}) of a known amount of purified PCR product obtained from genomic *E. coli* DNA by using the bacterial 16S rDNA-specific primers 8F/1392R. *Sulfolobus solfataricus* genomic DNA was used as negative control. The R^2 value for the standard curve was 0.99 and slope value was –3.48, giving an estimated amplification efficiency of 94%. *Aquificales* 16S rRNA genes were quantified using the *Aquificales*-specific primers as described (aqfx540modF/Aqui1197R) and under the same thermal conditions. Quantification standard consisted of a dilution series (between 10 and 1×10^6 copies μl^{-1}) of a known amount of purified PCR product obtained from DNA extracted from the *Aquificales*-positive sites G1, G13, Kam36 and Kam40. Genomic DNA from *Sulfolobus solfataricus* and *E. coli* were used as negative controls. The R^2 value for the standard curve used was 0.99 and slope value was –3.26, giving an estimated amplification efficiency of 101%.

All qPCR experiments were conducted in a StepOnePlus real-time PCR system (Applied Biosystems Inc., Foster City, CA, USA) with SYBRGreenI as fluorescent dye. To confirm product specificity, melting curve analyses were performed after each run for all experiments and each qPCR setup contained samples and standard series in triplicates and negative and blank in duplicates. All DNA concentrations were measured with a Nanodrop ND-1000 (Nanodrop

Technologies, Thermo Fisher Scientific Inc., Wilmington, DE, USA) and a Cary 300 BIO UV-visible spectrophotometer (Varian Inc., Palo Alto, CA, USA).

Archaeal and 'universal' 16S rDNA clone libraries
For taxonomic profiling of a hot spring community where *Korarchaeota* are present, we constructed an archaeal and a 'universal' 16S rRNA gene clone library from site IceG13 (84 °C, pH 5.5). The archaeal clone library (126 clones) was made with primers 20F/958R, and all clone inserts were sequenced (as described before). The 'universal' 16S rRNA gene clone library (118 clones) was constructed using primers 515F/1408R (Takai and Sako, 1999). The following PCR program was used for both primer sets: 95 °C for 5 min, then 25 cycles at 95 °C for 45 s, 55 °C for 45 s and 72 °C for 60 s, followed by 72 °C for 10 min and cooling (4 °C).

Phylogenetic analyses
Before phylogenetic analyses, 301 sequences from this study and 67 putative *Korarchaeota* 16S rDNA sequences from GenBank (per May, 2009), were subjected to chimaera check (Cole *et al.*, 2007), Bellerophon (Huber *et al.*, 2004) and sequence identity (ID) matrix analyses (BioEdit Sequence Alignment Editor v7.0.5.5) (Hall, 1999). Whereas no chimaeras were identified among our sequences, four sequences from public databases were suspected to be chimaeras (DQ228630, AB095127, AY862080 and AY861978). In all, 186 sequences shared 100% identity with *Korarchaeota* sequences originating from the same geographic location (181 of our sequences and five from GenBank: AB293207, AB293211, AB293208, DQ465911 and FN178533). Along with three sequences deposited in GenBank as *Korarchaeota* (EF100620, EF100633 and AY861978) that did not affiliate with the *Korarchaeota* kingdom in our analyses, the chimaeric and the identical sequences were removed from our phylogenetic analyses. In total, 175 sequences were aligned in the ARB program (v07.12.07org) (Ludwig *et al.*, 2004) against the ARB-compatible SILVA Release96 (Pruesse *et al.*, 2007). Phylogenetic relationships were calculated applying a 50% similarity filter as well as archaeal positional variability filter (pos_var_Archaea_96) in the ARB packages, using *E. coli* positions 264–1228 and excluding columns with high variability, thus leaving 911 valid columns. The alignment was subjected to neighbour joining analyses using Felsenstein correction, maximum likelihood (PhyML (Guindon and Gascuel, 2003) and RaxML (Stamatakis *et al.*, 2005) applying the Hasegawa-Kishono-Yano (HKY) or HKY85 nucleotide substitution model, respectively) and maximum parsimony (Phylip DNAPARS (Felsenstein, 1989)) analyses using the software integrated in the ARB packages. However, 30 sequences did not contain complete information in the selected region and these were added to the trees applying the parsimony tool in the ARB program, using 581 (29 sequences) or 441 (1 sequence) valid columns for calculation (Supplementary Figure 1 and Table 1). Five short sequences were completely removed as their phylogenetic affiliation could not be resolved (AF311360, AF411237, FN178541, FN178554 and FN178584). For tree simplicity, sequences from our study with $>99\%$ similarity and originating from the same geographic location have subsequently been removed from the tree in Figure 1. A full uncollapsed tree is provided in Supplementary Figure 1.

16S rDNA sequences of *Aquificales* were aligned using the BioEdit Sequence Alignment Editor v7.0.5.5 (Hall, 1999). Neighbour joining analyses were performed using BioEdit software on 575 homologous positions and bootstrapping was calculated based on 1000 resamplings.

Measurement of abiotic factors
At 2 h after sampling, liquid from three Icelandic springs, all *Korarchaeota* positive, were harvested by centrifugation ($10\,000 \times g$ for 5 min) and nitric acid was added to the supernatant (2% final concentration). Geochemical composition of liquids was analysed using Finnigan Element2 inductively coupled plasma mass spectrometer (Thermo Fisher Scientific Inc., Waltham, MA, USA) and Metrohm ion chromatograph (Metrohm Schwerz AG, Zofingen, Switzerland) at Centre for Element and Isotope Analysis at the University of Bergen (Table 3).

Measurement of total organic carbon (TOC)
The content of total organic carbon (TOC) of each sampling site was measured in a Multi EA 4000 carbon analyser using software MultiWin v4.07 (AnalytikJena, Jena, Germany). Samples were combusted at 550 °C in an oxygen-rich atmosphere. At this temperature, the organic carbon in the sample is converted to carbon dioxide and the concentration measured (Dean, 1974). Known concentration of calcium carbonate was used as standard. Frozen sample material was thawed on ice and dried at 100 °C overnight and TOC analysis was performed on the dried material.

Nucleotide sequence accession numbers
The 120 unique *Korarchaeota* 16S rRNA gene sequences reported in this study have been deposited in GenBank with accession numbers GQ228689–GQ228808. The IceG13 *Aquificales* 16S rRNA gene sequence identified is deposited in GenBank with the accession number GQ228809.

Results

Detection and cloning of 16S rRNA genes of Korarchaeota
We analysed 19 terrestrial hot springs from Kamchatka and Iceland for the presence of

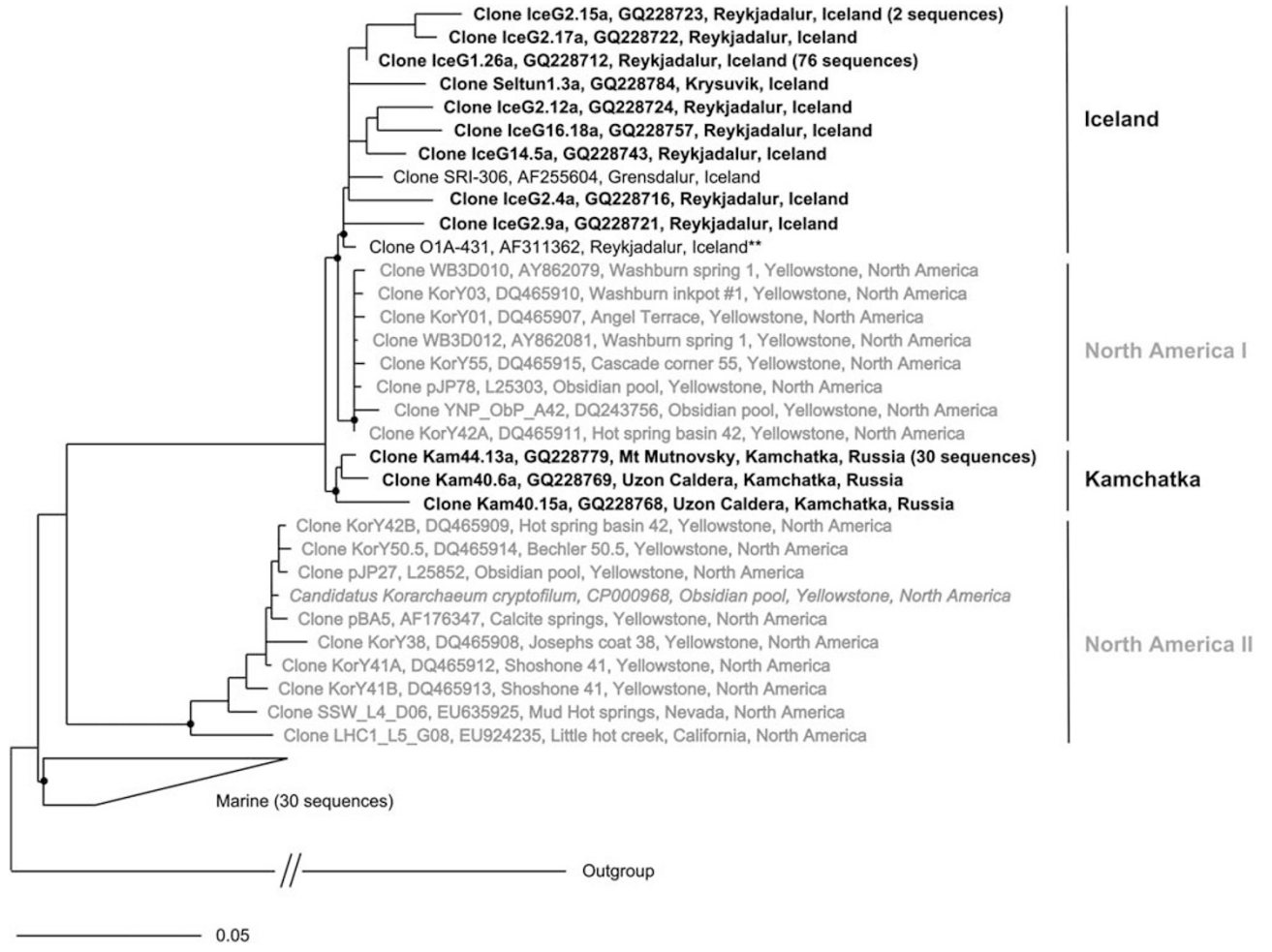

Figure 1 *Korarchaeota* phylogeny based on 16S rDNA information. The maximum likelihood tree (RaxML with Hasegawa–Kishono–Yano (HKY)85 substitutions) was determined by analysis of 911 valid columns aligned and calculated in the ARB package. A total of 30 sequences were subsequently added using parsimony tool with 581 valid columns or 441 valid columns. As outgroup the following sequences were used: AB109559, D83259, X99554, Y07784, AF191225, AB087499, M21087, X14835, AF255605, AB293224, AF419633, AB063630, D85506, X90484, X89852, X90478, D85520, X90480, D85507 and AB010957. Black dots indicate a branching point supported by all four phylogenetic calculations (neighbour joining (NJ), PhyML, RaxML and Phylip DNAPARS). Sequences obtained in this study that share >99% of the sequence similarity and originate from the same geographic location have been removed after the phylogenetic calculation. The number of removed sequences is given in brackets after their reference sequence. Bold text indicates sequences obtained in this study. Scale bar: 5 substitutions per 100 nucleotides. The double asterisk indicates that the sequence has been added using 441 valid columns. A non-collapsed tree is provided in Supplementary Figure 1.

Korarchaeota using 16S rRNA-specific primers Kor236F/Kor1236R (Auchtung *et al.*, 2006) in PCR. The sample materials were collected from sites with temperatures $\geqslant 70\,^{\circ}C$ and with pH from 2.5 to 6.5 (Table 1). The organic carbon contents were in the range of 0.01–0.46% of the wet weight (0.06–1.85% of dry weight). The acidic mud samples showed generally lower TOC values than the filaments and microbial mat samples (Table 1).

We found 12 out of 19 hot springs to be positive for *Korarchaeota* (8 out of 12 in Iceland and 4 out of 7 in Kamchatka). Whereas *Korarchaeota* 16S rRNA genes were detected in DNA from all hot spring filaments and one of the microbial mats (pH 5.5–6.5), none of the acidic mud samples (pH 2.5–4.5) tested positive for *Korarchaeota,* although specific PCR products were obtained with 'general' Archaea 16S rDNA primers. From each of the 12 positive sites, a *Korarchaeota* 16S rDNA clone library was made and a total of 301 randomly chosen clones were subjected to sequencing. Followed by sequence ID matrix analyses, 16S rDNA sequences sharing 100% similarity, as well as poor-quality sequences, were removed from the analyses. This resulted in 33 unique Kamchatkan *Korarchaeota* 16S rRNA gene sequences and 87 Icelandic. It is noteworthy that the similarity between these new sequences and that of *Candidatus Korarchaeum cryptofilum* (Elkins *et al.*, 2008) is $\leqslant 93.5\%$, indicating a distant relationship.

Table 2 Enumeration of *Korarchaeota,* Archaea, Bacteria and *Aquificales* 16S rDNA copies from hot springs in Kamchatka and Iceland using qPCR

Site	*Korarchaeota 16S rDNA copies ng^{-1} DNA*	*Archaea 16S rDNA copies ng^{-1} DNA*	*Bacteria 16S rDNA copies ng^{-1} DNA*	*Aquificales 16S rDNA copies ng^{-1} DNA*	*Korarchaeota of total*[a] *(%)*	*Archaea of total*[a] *(%)*	*Aquificales of total*[a] *(%)*	*Korarchaeota of Archaea*[b] *(%)*
IceG1	$23.5 \times 10^2 \pm 138$	$7.3 \times 10^4 \pm 23$	$25.7 \times 10^4 \pm 1209$	$74.8 \times 10^4 \pm 902$	0.2	6.7	69.2	3.2
IceG2	$1.2 \times 10^2 \pm 29$	$0.8 \times 10^4 \pm 821$	$12.9 \times 10^4 \pm 4687$	$14.0 \times 10^4 \pm 78$	0.04	2.8	50.4	1.5
IceG13[c]	$8.6 \times 10^2 \pm 8$	$1.3 \times 10^4 \pm 845$	$17.6 \times 10^4 \pm 11\,265$	$54.1 \times 10^4 \pm 1100$	0.4	6.4	71.1	6.8
IceG14	$3.6 \times 10^2 \pm 58$	$4.8 \times 10^4 \pm 1094$	$1.21 \times 10^4 \pm 1638$	$7.11 \times 10^4 \pm 890$	0.3	36.4	54.1	0.8
IceG16	$0.52 \times 10^2 \pm 8$	$5.7 \times 10^4 \pm 9030$	$0.38 \times 10^4 \pm 2790$	$8.13 \times 10^4 \pm 1072$	0.04	40.0	57.2	0.1
Seltun1	$0.8 \times 10^2 \pm 7$	$9.9 \times 10^4 \pm 1100$	$0.42 \times 10^4 \pm 294$	$5.68 \times 10^4 \pm 638$	0.05	61.8	35.5	0.1
Seltun3	$0.1 \times 10^2 \pm 4$	$5.4 \times 10^4 \pm 598$	$7.68 \times 10^4 \pm 4589$	$37.4 \times 10^4 \pm 6074$	0.002	10.7	74.1	0.02
Kam36	$7.8 \times 10^2 \pm 132$	$3.4 \times 10^4 \pm 1119$	$3.29 \times 10^4 \pm 2100$	$11.1 \times 10^4 \pm 866$	0.4	18.9	62.2	2.3
Kam38	$0.05 \times 10^2 \pm 0.05$	$11.0 \times 10^4 \pm 611$	>100	$15.0 \times 10^4 \pm 1686$	0.002	42.3	57.7	0.005
Kam40	$0.78 \times 10^2 \pm 5$	$11.7 \times 10^4 \pm 9768$	$1.35 \times 10^4 \pm 816$	$0.99 \times 10^4 \pm 323$	0.6	82.8	7.0	0.7
Kam44	$6.6 \times 10^2 \pm 30$	$3.4 \times 10^4 \pm 72$	$23.2 \times 10^4 \pm 2306$	$5.12 \times 10^4 \pm 559$	0.2	10.7	16.1	2.0

[a]The total number is the sum of Bacteria, Archaea and *Aquificales* 16S rDNA copies ng^{-1} DNA.
[b]Percentage of korarchaeal 16S rDNA copies ng^{-1} DNA of total Archaea 16S rDNA copies ng^{-1} DNA.
[c]Quantitative PCR (qPCR) data for IceG13A are not available.

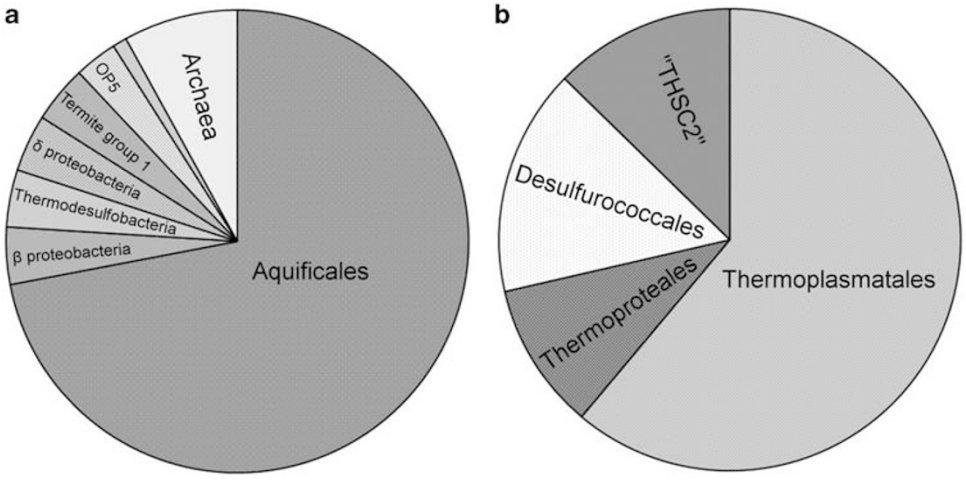

Figure 2 16S rDNA community profile from the *Korarchaeota*-positive Icelandic hot spring IceG13. (**a**) Clone library (118 sequences) made with 'universal' primers (515F/1408R). The 16S rDNA sequences distributed as follows: 69% *Aquificales*, 8% Archaea, 5% δ-*proteobacteria,* 5% *Thermoplasmatales*, 4% *candidate division OP5*, 4% β-*proteobacteria*, 4% *Termite group 1* and 1% *Nitrospirae*. (**b**) Clone library (126 sequences) made with Archaea-specific primers (20F/958R). *Thermoplasmatales* constituted 61% of the sequences, *Desulfurococcales* 16%, *Terrestrial hot spring cluster 2* (*THSC2*) 13% and *Thermoproteales* 10%.

Quantitative estimates of Korarchaeaota, Aquificales, *Bacteria and Archaea in terrestrial hot springs*

Further insights into *Korarchaeota*-positive communities were obtained by enumeration of 16S rRNA genes. To establish the relative contribution of *Korarchaeota* to the total microbial population, Bacteria, Archaea and *Korarchaeota* 16S rRNA genes were quantified using qPCR (Table 2). Sequences of *Aquificales* are not well covered by the bacterial primers (1214 out of 1282 *Aquificales* sequences have $\geqslant 2$ mismatches to the bacterial forward primer used in this study), but often represent the dominating group in terrestrial hot springs. We have therefore used an additional primer set that covers specifically this phylum. The copy number of *Korarchaeota* 16S rRNA genes ranged between 5.4×10^0 and 2.4×10^3 copies per ng DNA and thus constituted $<1\%$ of the total community (0.002–0.6%) and between 0.02% and 6.8% of the archaeal population. The qPCR enumeration of archaeal 16S rDNA showed that the relative amount of Archaea in the microbial communities was between 2.8% and 82.8%. In all *Korarchaeota*-positive hot springs, *Aquificales* 16S rDNA were detected. With DNA ranging between 9.9×10^3 and 7.5×10^5 copy numbers per ng DNA, they made up between 18% and 100% of the bacterial population. The relative amounts of Archaea, Bacteria and *Aquificales* were well supported by the number of 16S rRNA genes-containing genomic clones found within large-insert meta-genomic libraries (not shown) and by distribution in the clone libraries (Table 2 and Figure 2).

Table 3 Chemical composition of spring fluids from three Icelandic hot springs harbouring *Korarchaeota*

Element ($\mu g\,l^{-1}$)	*Site IceG1*	*Site IceG13*	*Site Seltun1*
Lithium (Li)	15.8	0.2	5.7
Boron (B)	24.3	2.4	14.9
Sodium (Na)	1 969 096.1	8712.5	20 756.8
Magnesium (Mg)	11 757.1	9442.6	13 134.0
Aluminum (Al)	1056.6	12 345.0	88.0
Potassium (K)	ND	1557.0	ND
Calcium (Ca)	244 323.5	18 812.0	46 874.3
Manganese (Mn)	4701.1	46.0	73.5
Iron (Fe)	3448.3	626.0	83.0
Cobalt (Co)	2.5	0.1	0.2
Copper (Cu)	ND	35.0	ND
Arsenic (As)	8.3	1.1	0.2
Molybdenum (Mo)	2.6	1.6	0.3
Barium (Ba)	73.5	3.0	2.7
Anion ($mg\,l^{-1}$)			
Fluoride (F^-)	33.8	0.2	0.6
Chloride (Cl^-)	867.3	6.7	9.9
Sulphate (SO_4^{2-})	204 322.3	111.0	286.5
Nitrate (NO_3)	ND	Trace	ND

Abbreviation: ND, not determined.

16S rRNA phylogeny of Korarchaeota

The collection of *Korarchaeota* 16S rRNA genes used for phylogenetic analyses consisted of 170 sequences (50 from public databases and 120 from this study; Supplementary Table 1). We identified a separation between terrestrial and marine-retrieved sequences (Figure 1 and Supplementary Figure 1). Terrestrial sequences formed four sub-clusters (Iceland, North America I and II, and Kamchatka). The Icelandic cluster contained all Icelandic sequences from this study as well as two sequences previously obtained from Icelandic terrestrial hot springs (Hjorleifsdottir *et al.* 2001; Skirnisdottir *et al.*, 2000). The North America cluster I was restricted to sequences retrieved from Yellowstone National Park, whereas the North America cluster II included sequences from Yellowstone National Park as well as two sequences from Californian hot springs. The fourth and last terrestrial cluster consisted of sequences obtained from Kamchatka. The marine sequences all grouped together in a large cluster containing exclusively sequences from the marine environment with one exception, that is, the sequence from clone pUWA43 retrieved from an acidic terrestrial Japanese hot spring (Takai and Sako, 1999). Several smaller clusters within the marine group were observed without any apparent geographic distribution pattern. The main topology with four terrestrial clusters and a distinct separation between marine and terrestrial sequences in the phylogenetic tree were supported by all analyses performed (neighbour joining, PhyML, RaxML and Phylip DNAPARS) as indicated by black dots in Figure 1. In none of the analyses was a terrestrial sequence placed within the marine cluster or vice versa (except the mentioned pUWA43).

Abiotic and biotic characterization of Korarchaeota*-positive environments*

To provide more information on the environments where *Korarchaeota* thrive, we performed chemical analyses of three Icelandic hot springs (IceG1, IceG13 and Seltun1) in which we had detected *Korarchaeota* 16S rRNA genes (Table 3). The water chemistry was strikingly different: Spring IceG1 resembled sea water with respect to measured levels of sodium, chloride, magnesium, sulphate and calcium and was also high in ferrous iron. Spring IceG13 had very high content of aluminium but lower levels of iron than IceG1. Seltun1 featured low levels of both aluminium and iron.

To characterize the microbial community in IceG13, we constructed an archaeal and a 'universal' 16S rRNA gene library and sorted the sequences into major phylogenetic groups using the ARB program and Greengenes tools (DeSantis *et al.*, 2006). The 'universal' clone library (Figure 2a) contained 92% sequences of Bacteria and 8% of Archaea. It was dominated by *Aquificales* with 69% of the clones (81 of 118 clones) with one single sequence representing 90% of these. This sequence was >99% similar to the 16S rRNA gene of the newly isolated *Sulfurihydrogenibium kristjanssonii* cultivated from a hot spring at Hveragerdi (Flores *et al.*, 2008). This sequence affiliates with the *Hydrogenothermaceae* family. Interestingly, the *Aquificales* previously shown to dominate other *Korarchaeota*-positive terrestrial hot springs also belonged to this family: clone pBB (Reysenbach *et al.*, 2000), clone OPB13 (Hugenholtz *et al.*, 1998) and clone SRI-240 (Skirnisdottir *et al.*, 2000). In addition, phylogenetic analysis showed that within the *Hydrogenothermaceae* family, all these four uncultured *Aquificales* affiliate with the *Sulfurihydrogenibium* genus (Figure 3). The possible coexistence of *Korarchaeota* and *Aquificales* inspired us to test all 19 investigated hot springs for the presence of *Aquificales,* using specific primers. All 12 *Korarchaeota*-positive springs were positive for *Aquificales* (Table 1). In addition, two sites, IceG3 and IceR11, were positive for *Aquificales*, but negative for *Korarchaeota.*

The remaining bacterial sequences in the 16S rDNA clone library of IceG13 affiliated with *β*- and *δ-Proteobacteria* (4% and 5%, respectively), *Thermodesulfobacteria* (5%), *Nitrospirae* (1%), *Termite gr.1* (4%) and *Candidate division OP5* (4%) (Figure 2a).

In the archaeal 16S rRNA gene library (Figure 2b) made in parallel to the 'universal' library, 61% of the clones were affiliated with *Thermoplasmatales,* 16% with *Desulfurococcales,* 13% with *Terrestrial hot spring cluster 2* and 10% with *Thermoproteales*. Although *Korarchaeota* sequences were amplified from this spring with specific primers, they were not detected among the 244 sequenced clones from the two 16S rRNA clone libraries, nor in cDNA made from this Icelandic site (not shown). However, primer 958R, which is used to make the archaeal

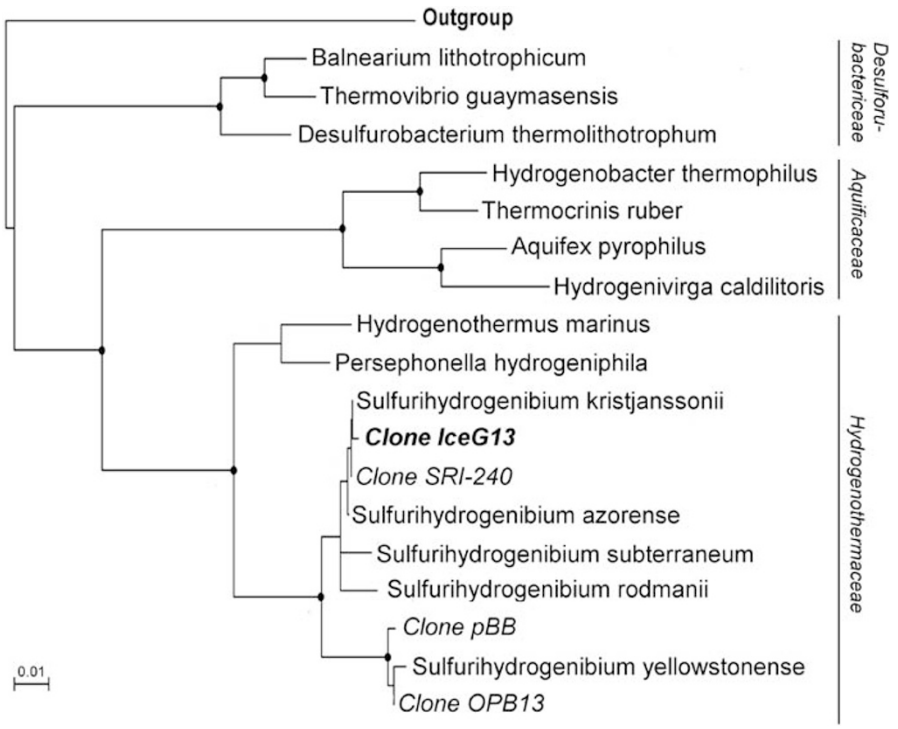

Figure 3 16S rDNA phylogram of *Aquificales* dominating microbial communities in *Korarchaeota*-positive terrestrial hot springs. *Aquificales* found together with *Korarchaeota* are shown in italic (Clone pBB, AF113542; clone OBP13, AF027098; clone IceG13 from this study (bold italic), GQ228809; and clone SRI-240, AF255597). Cultured *Aquificales* included in phylogenetic analyses are retrieved from GenBank (AJ001049, AB105049, AY268937, M83548, AB120294, Z30189, AJ005640, AB086419, AJ292525, AY686713, AB071324, AF528192, AM778960 and AM259495). The tree was rooted with *Thermotoga maritima* (M21774). Neighbour joining (NJ) analysis was carried out using 575 homologous nucleotides with 1000 resamplings. Scale bar: 1 substitution per 100 nucleotides.

clone library, has one mismatch with the *Korarchaeota* group, and primer 20F, also used for the archaeal library, can have several mismatches.

Discussion

In our survey of 19 terrestrial hot springs from Iceland and Kamchatka with temperatures between 70 °C and 97 °C and pH 2.5–6.5, we detected and enumerated *Korarchaeota* 16S rDNA in 12 springs. The *Korarchaeota*-positive hot springs had temperatures between 70 °C and 84 °C and pH between 5.5 and 6.5 (Table 1). This temperature regime is supported by all, but one, other study detecting *Korarchaeota* 16S rDNA. The exception was detection in organic material being ~2 °C on an abyssal hill of the East Pacific Rise, a site where a subseafloor hydrothermal habitat beneath the ridge flank was suggested as a source of hyperthermophilic organisms (Ehrhardt *et al*., 2007b). Concerning the pH of the *Korarchaeota*-positive sites, there are two reports indicating that *Korarchaeota* are able to sustain both a more alkaline and a more acidic environment than the (sub-)neutral pH found in our studies: a Japanese hot spring of Mt Unzen with pH 2.8 at 128 °C (Takai and Sako, 1999), and a smectite cone off Iceland with pH 10 at 72 °C (Marteinsson *et al*., 2001b).

Three types of sampling material were analysed: filaments, microbial mats and mud. Interestingly, all filaments and one of the microbial mats were positive for *Korarchaeota* but none of the mud samples (Table 1). This is in agreement with other studies, in which *Korarchaeota* sequences have predominantly been retrieved from similar environments, such as the filaments and mats. It has been suggested that *Korarchaeota* receive some essential nutrients from other mat-forming organisms (Elkins *et al*., 2008). This could be one reason why we only detected *Korarchaeota* 16S rRNA genes in filament and mat samples that were mostly richer in biomass (see TOC in Table 1) and not in the bubbling (acidic) mud samples. Our study further indicates that *Korarchaeota* might specifically co-occur with *Aquificales* or might live in syntrophy with other organisms that are dependent on *Aquificales*, as all 12 *Korarchaeota*-positive hot springs were also positive for *Aquificales*.

Based on the 'universal' 16S rDNA clone library from the *Korarchaeota*-positive filaments in spring IceG13, the community was clearly dominated by bacteria (Figure 2a and Table 2) with the most abundant bacterial sequence belonging to a single phylotype within *Aquificales,* the newly isolated

thermophilic hydrogen and sulphur oxidizer *Sulfurihydrogenibium kristjanssonii* cultivated from a hot spring in Hveragerdi (Flores *et al.*, 2008). The observation that *Korarchaeota* coexist with *Aquificales* has been reported previously for Calcite Springs (83 °C, pH 7.6) (Reysenbach *et al.*, 2000) and Obsidian Pool (80 °C, pH 6.5) (Hugenholtz *et al.*, 1998; Spear *et al.*, 2005), both Yellowstone, and for the submarine smectite cones off the coast of Iceland (Marteinsson *et al.*, 2001b). A closer analysis into the phylogeny revealed that all the *Aquificales* that dominated these four microbial communities belonged to the *Hydrogenothermaceae* family and could be further designated to the *Sulfurihydrogenibium* genus (Figure 3), a genus well known for appearing as filamentous biomass in hot springs. It was therefore suggested that the *Korarchaeota* and the *Aquificales* have overlapping niches and that some commonality may exist in their growth requirements, such as pH near neutrality, iron-rich environments and overlapping temperatures for growth (Reysenbach *et al.*, 2000). However, with genome information from *Candidatus Korarchaeum cryptofilum,* one might as well speculate that the coexistence with *Aquificales* is needed to remove the hydrogen produced by *Korarchaeota* during fermentation of peptides and amino acids. Whatever the reason for their co-occurrence, we could not find any apparent correlation between the absolute numbers of *Aquificales* and *Korarchaeota* (Table 2).

As stated previously, the community of the analysed spring harbours few Archaea compared with bacteria, but of those Archaea present, the vast majority (61%) is *Thermoplasmatales*-related species (Figure 2b), known to thrive under such subneutral conditions.

Quantitative PCR measurements indicated that *Korarchaeota* constitute only a minor fraction of the entire microbial community in the analysed *Korarchaeota*-positive hot springs (Table 2). Even where Archaea represented a large fraction of the entire community (up to 83%), the *Korarchaeota* were found to constitute $<1\%$ of the total community. These relatively low numbers support the information that has been deduced from the *Candidatus Korarchaeum cryptofilum* genome. It is not likely that a high number of organisms can be sustained with a lifestyle in which energy is gained by fermentation of peptides and in which the organism depend on others to produce essential vitamins, cofactors, purines and possibly the removal of hydrogen.

Despite being present in low amounts, there seemed to be significant spring-to-spring variation in the number of *Korarchaeota* 16S rDNA, although the springs were located only a few meters apart. This was the case for the Icelandic springs IceG1 and IceG2 (20-fold difference) and between Kam36 and Kam38 (>100-fold difference).

After search in databases, an up-to-date collection of available *Korarchaeota* 16S rRNA genes was made, resulting in 50 database-retrieved sequences and 120 sequences from this study (Supplementary Table 1). All phylogenetic analyses performed showed a clear separation between terrestrial and marine sequences. Never did a terrestrial sequence affiliate within a marine cluster or vice versa (Figure 1), the only exception being the terrestrial sequence pUWA43, placed inside the marine cluster. However, in this case the source fluid might be a mixture of magmatic volatiles and sea water and the sequence could therefore very well be of marine origin (K Takai, personal communication). All terrestrial sequences lie within four sub-clusters, showing a geographic pattern.

Our Icelandic sequences clustered together with two previously reported terrestrial Icelandic *Korarchaeota* sequences (Hjorleifsdottir *et al.*, 2001; Skirnisdottir *et al.*, 2000). There are two other published *Korarchaeota* sequences from Iceland, but these were too short for their phylogenetic affiliation to be resolved (AF311360 and AF411237). However, AF311360 shares the two positions unique for all Icelandic sequences (AF411237 has no information in this region) and affiliates within the Icelandic cluster using maximum likelihood and maximum parsimony, but cannot be resolved using neighbour joining analysis. The *Korarchaeota* sequences from Kamchatka formed a novel and separate group within the terrestrial branch. Previously detected *Korarchaeota* 16S rRNA genes from Yellowstone and California form two separate clusters. This finding supports a link between geographic distribution and phylogenetic relation within the terrestrial *Korarchaeota*.

However, the marine *Korarchaeota* do not seem to show a geographic pattern. An explanation could be that terrestrial *Korarchaeota* live in habitats that are more separated than those of the marine counterparts. It has been argued by (Fenchel, 2003) that cell transfer from one spring to another is a very unlikely event because of the fact that terrestrial hot springs around the globe are small in size and geographically far apart. In contrast, the marine environment is more homogenous and dispersal of cells from one submarine geothermal area to another seems far more likely. It should also be noticed that all but one of the marine sequences stem from the Pacific Ocean. Whether the terrestrial sequence affiliations in this study show a true geographic distribution or are an artefact because of the relatively low number of 16S rDNA sequences available from a limited number of habitats is not known.

Our results further emphasize the surprisingly low diversity within the terrestrial *Korarchaeota*. Indeed, it could be argued that several sequences within each cluster belong to the same species and there are several cases in which the same applies for sequences from different clusters (using the 97% species criteria and approximately 1000 bp). Considering that we did not use a polymerase with 3′–5′ exonuclease activity, several of the differences that are observed between sequences within each cluster

are likely to be caused by PCR errors and/or sequencing errors, making the sequence homogeneity even more stunning. However, a few highly conserved positions seem to be unique for each geographic location. To unravel the reason behind this intriguing observation, more studies and genomic information are needed.

In analysing the geological settings in the *Korarchaeota*-positive springs in our and other studies, it becomes clear that *Korarchaeota* thrive in both basalt-hosted (Hveragerdi and Krysuvik, Iceland) and rhyolite-hosted (Obsidian pool and Calcite Springs, Yellowstone) hot springs, two bedrocks that are different with regard to both their chemical compositions and their dissolving rates when in contact with water. Furthermore, the water chemistry analyses for three of the *Korarchaeota*-positive Icelandic springs clearly show that the habitats are significantly different despite all three springs being located in basaltic lava bedrock (Table 3). The IceG1 hot spring is located at the very beginning of the Reykjadalur, which is almost at sea level. The concentration of sodium, chloride, magnesium, sulphate and calcium in IceG1 indicate that sea water has been mixed into the hydrothermal water and/or cold groundwater at some point before reaching the surface. Still, the iron content reflects influence of thermal activity. The IceG13 spring is located further up in Reykjadalur, >200 altitude meters higher and its water chemistry seems not to be influenced by sea water. The very high content of aluminium might reflect decomposition of basaltic lava at a greater extent compared with the other springs that were analysed. Seltun1 spring, which is 160 m in height in Krysuvik, is not affected by sea water, and contains low levels of aluminium and iron, indicating a very different local environment when compared with IceG1 and IceG13. Our study and studies from the Obsidian pool and Calcite springs of Yellowstone show *Korarchaeota* to thrive together with *Aquificales* in sub-neutral springs above 70 °C, and a comparison of the water chemistry of these two springs of Yellowstone (Ball *et al*., 1998) with the IceG1, IceG13 and Seltun1 shows *Korarchaeota* to thrive in diverse spring fluid chemistry in which the levels of manganese, aluminium and iron seem not to influence neither presence nor abundance of *Korarchaeota*.

In conclusion, our study increases the understanding of the diversity, distribution and abundance of *Korarchaeota* in hot terrestrial environments. The qPCR techniques established in this study for *Korarchaeota* and *Aquificales* will be useful in dissecting environments in which *Korarchaeota* are naturally more enriched. Those environments could be used as a starting point for cultivation attempts, physiological characterizations and metagenomic or genomic studies that will help to further elucidate the phylogeny and metabolism of this ancient group of hyperthermophilic Archaea.

Acknowledgements

We thank the Microbial Observatory of Kamchatka, especially Lisa Bonch-Osmolovskaya, for the possibility to attend the workshop in 2005: 'Biogeochemistry and molecular biology of extremophiles' and to sample from hot springs on the Kamchatka Peninsula (LJR). We also thank Jakob Kristjansson for the possibility to sample in Hveragerdi in 2005, 2006 and 2007. We would acknowledge Tryggvi Thordarson for access to laboratory facilities at the University of Iceland Research Station in Hveragerdi. We thank Tim Urich for help during sampling in Iceland 2006, and Ingunn Thorseth for help with geochemistry data. This work was supported by the Norwegian Research Council, Grant 172206 to CS and LJR.

References

Auchtung TA, Takacs-Vesbach CD, Cavanaugh CM. (2006). 16S rRNA phylogenetic investigation of the candidate division 'Korarchaeota'. *Appl Environ Microbiol* **72**: 5077–5082.

Ball, Nordstrom DK, Jenne EA, DV V. (1998). Chemical analyses of hot springs, pools, geysers, and surface waters from Yellowstone National Park, Wyoming, and vicinity, 1974–1975. *USGS Open-File Report* **182**: 1–45.

Barns SM, Delwiche CF, Palmer JD, Pace NR. (1996). Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proc Natl Acad Sci USA* **93**: 9188–9193.

Barns SM, Fundyga RE, Jeffries MW, Pace NR. (1994). Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment. *Proc Natl Acad Sci USA* **91**: 1609–1613.

Brochier-Armanet C, Boussau B, Gribaldo S, Forterre P. (2008). Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat Rev Microbiol* **6**: 245–252.

Burggraf S, Heyder P, Eis N. (1997). A pivotal Archaea group. *Nature* **385**: 780.

Cavicchioli R. (2007). *Archaea: Cellular and Molecular Biology*. ASM Press: Washington DC, USA, 556.

Cole JR, Chai B, Farris RJ, Wang Q, Kulam-Syed-Mohideen AS, McGarrell DM *et al.* (2007). The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data. *Nucleic Acids Res* **35**(Database issue): D169–D172.

Dean Jr WE. (1974). Determination of carbonate and organic matter in calcareous sediments and sedimentary rocks by loss on ignition: comparison with other methods. *J Sediment Petrol* **44**: 242–248.

DeLong EF. (1992). Archaea in coastal marine environments. *Proc Natl Acad Sci USA* **89**: 5685–5689.

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K *et al.* (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* **72**: 5069–5072.

Ehrhardt CJ, Haymon RM, Lamontagne MG, Holden PA. (2007a). Evidence for hydrothermal Archaea within the basaltic flanks of the East Pacific Rise. *Environ Microbiol* **9**: 900–912.

Ehrhardt CJ, Haymon RM, Lamontagne MG, Holden PA. (2007b). Evidence for hydrothermal Archaea within the basaltic flanks of the East Pacific Rise. *Environ Microbiol* **9**: 900–912.

Einen J, Thorseth IH, Ovreas L (2008). Enumeration of Archaea and Bacteria in seafloor basalt using real-time quantitative PCR and fluorescense microscopy. *FEMS Mikrobiol Lett* **282**: 182–187.

Elkins JG, Podar M, Graham DE, Makarova KS, Wolf Y, Randau L *et al.* (2008). A korarchaeal genome reveals insights into the evolution of the Archaea. *Proc Natl Acad Sci USA* **105**: 8102–8107.

Felsenstein J. (1989). PHYLIP—phylogeny inference package (version 3.2). *Cladistics* **5**: 164–166.

Fenchel T. (2003). Microbiology. Biogeography for bacteria. *Science* **301**: 925–926.

Flores GE, Liu Y, Ferrera I, Beveridge TJ, Reysenbach AL. (2008). Sulfurihydrogenibium kristjanssonii sp. nov., a hydrogen- and sulfur-oxidizing thermophile isolated from a terrestrial Icelandic hot spring. *Int J Syst Evol Microbiol* **58**(Part 5): 1153–1158.

Garrett R, Klenk HK. (2006). *Archaea: Evolution, Physiology, and Molecular Biology*. Blackwell Publishing: Oxford, UK.

Guindon S, Gascuel O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**: 696–704.

Hall T. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* **41**: 95–98.

Hjorleifsdottir S, Skirnisdottir S, Hreggvidsson GO, Holst O, Kristjansson JK. (2001). Species composition of cultivated and noncultivated bacteria from short filaments in an Icelandic hot spring at 88 degrees C. *Microb Ecol* **42**: 117–125.

Huber H, Hohn MJ, Rachel R, Fuchs T, Wimmer VC, Stetter KO. (2002). A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* **417**: 63–67.

Huber T, Faulkner G, Hugenholtz P. (2004). Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* **20**: 2317–2319.

Hugenholtz P, Pitulle C, Hershberger KL, Pace NR. (1998). Novel division level bacterial diversity in a Yellowstone hot spring. *J Bacteriol* **180**: 366–376.

Leininger S, Urich T, Schloter M, Schwark L, Qi J, Nicol GW *et al.* (2006). Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature* **442**: 806–809.

Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar *et al.* (2004). ARB: a software environment for sequence data. *Nucleic Acids Res* **32**: 1363–1371.

Marteinsson VT, Hauksdottir S, Hobel CF, Kristmannsdottir H, Hreggvidsson GO, Kristjansson JK. (2001a). Phylogenetic diversity analysis of subterranean hot springs in Iceland. *Appl Environ Microbiol* **67**: 4242–4248.

Marteinsson VT, Kristjansson JK, Kristmannsdottir H, Dahlkvist M, Saemundsson K, Hannington M *et al.* (2001b). Discovery and description of giant submarine smectite cones on the seafloor in Eyjafjordur, northern Iceland, and a novel thermal microbial habitat. *Appl Environ Microbiol* **67**: 827–833.

Nealson K. (2008). A Korarchaeote yields to genome sequencing. *Proc Natl Acad Sci USA* **105**: 8805–8806.

Nercessian O, Reysenbach AL, Prieur D, Jeanthon C. (2003). Archaeal diversity associated with *in situ* samplers deployed on hydrothermal vents on the East Pacific Rise (13 degrees N). *Environ Microbiol* **5**: 492–502.

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

Reysenbach AL, Banta A, Civello S, Daly J, Mitchel K, Lalonde S *et al.* (2005). *Aquificales* of Yellowstone National Park. In: Inskeep WP, McDermott TR (eds). *Geothermal Biology and Geochemistry in Yellowstone National Park, Bozeman, MT*. Thermal Biology Institute, Montana State University, Bozeman, MT, USA, 129–142.

Reysenbach AL, Ehringer M, Hershberger K. (2000). Microbial diversity at 83 degrees C in Calcite Springs, Yellowstone National Park: another environment where the Aquificales and 'Korarchaeota' coexist. *Extremophiles* **4**: 61–67.

Rogers KL, Amend JP. (2005). Archaeal diversity and geochemical energy yields in a geothermal well on Vulcano Island, Italy. *Geobiology* **3**: 319–332.

Rusch A, Amend JP. (2004). Order-specific 16S rRNA-targeted oligonucleotide probes for (hyper)thermophilic archaea and bacteria. *Extremophiles* **8**: 357–366.

Skirnisdottir S, Hreggvidsson GO, Hjorleifsdottir S, Marteinsson VT, Petursdottir SK, Holst O *et al.* (2000). Influence of sulfide and temperature on species composition and community structure of hot spring microbial mats. *Appl Environ Microbiol* **66**: 2835–2841.

Spear JR, Walker JJ, McCollom TM, Pace NR. (2005). Hydrogen and bioenergetics in the Yellowstone geothermal ecosystem. *Proc Natl Acad Sci USA* **102**: 2555–2560.

Stamatakis A, Ludwig T, Meier H. (2005). RAxML-III: a fast program for maximum likelihood-based inference of large phylogenetic trees. *Bioinformatics* **21**: 456–463.

Stetter KO. (2006). www.jgi.doe.gov/sequencing/why/CSP2006/Korarchaeota.html.

Takai K, Gamo T, Tsunogai U, Nakayama N, Hirayama H, Nealson KH *et al.* (2004). Geochemical and microbiological evidence for a hydrogen-based, hyperthermophilic subsurface lithoautotrophic microbial ecosystem (HyperSLiME) beneath an active deep-sea hydrothermal field. *Extremophiles* **8**: 269–282.

Takai K, Horikoshi K. (1999). Genetic diversity of archaea in deep-sea hydrothermal vent environments. *Genetics* **152**: 1285–1297.

Takai K, Moser DP, DeFlaun M, Onstott TC, Fredrickson JK. (2001). Archaeal diversity in waters from deep South African gold mines. *Appl Environ Microbiol* **67**: 5750–5760.

Takai K, Sako Y. (1999). A molecular view of archaeal diversity in marine and terrestrial hot water environments. *FEMS Microbiol Ecol* **28**: 177–188.

Teske A, Hinrichs KU, Edgcomb V, de Vera Gomez A, Kysela D, Sylva SP *et al.* (2002). Microbial diversity of hydrothermal sediments in the Guaymas Basin: evidence for anaerobic methanotrophic communities. *Appl Environ Microbiol* **68**: 1994–2007.

Treusch AH, Leininger S, Kletzin A, Schuster SC, Klenk HP, Schleper C. (2005). Novel genes for nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic crenarchaeota in nitrogen cycling. *Environ Microbiol* **7**: 1985–1995.

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)