The Journal of Microbiology, August 2009, p. 371-384 Vol. 47, No. 4
DOI 10.1007/s12275-008-0249-1
Copyright ⓒ 2009, The Microbiological Society of Korea

Molecular Analysis of Prokaryotic Diversity in the Deep Subsurface of the Former Homestake Gold Mine, South Dakota, USA

Gurdeep Rastogi[1#], Larry D. Stetler[2], Brent M. Peyton[3], and Rajesh K. Sani[1*]

[1]*Department of Chemical and Biological Engineering,* [2]*Department of Geology and Geological Engineering, South Dakota School of Mines and Technology, Rapid City, SD 57701, USA*
[3]*Department of Chemical and Biological Engineering, Montana State University, Bozeman, MT 59717, USA*
[#]*Present address: Department of Plant Pathology, University of California, Davis 95616, USA*

(Received October 6, 2008 / Accepted April 10, 2009)

A culture-independent molecular phylogenetic analysis was carried out to study the prokaryotic diversity in two soil samples collected from the subsurface (1.34 km depth) of the former Homestake gold mine, Lead, South Dakota, USA at two sites, the Ross shaft and number 6 Winze. Microbial community analyses were performed by cloning and sequencing of 16S rRNA genes retrieved directly from soil samples. Geochemical characterization of soils revealed high amount of toxic metals such as As, Cd, Co, Cr, Cu, Ni, Pb, Zn, and U at both the sites. Phylogenetic analyses showed that soil samples were predominantly composed of phylotypes related to phylum *Proteobacteria.* Other phyla detected in libraries were *Acidobacteria*, *Actinobacteria*, *Bacteroidetes*, *Chloroflexi*, *Chlorobi*, *Firmicutes*, *Gemmatimonadetes*, *Nitrospirae*, *Planctomycetes*, *Verrucomicrobia*, and candidate divisions OP10 and TM7. The majority (>95%) of the phylotypes retrieved in the libraries were most closely related to environmental sequences from yet-uncultured bacteria representing a hitherto unidentified diversity. The archaeal communities at both the sites exhibited lower diversity and were most closely affiliated to uncultivated species within the *Crenarchaeota.* Results showed the existence of diverse microbial populations in deep subsurface environment of the Homestake gold mine. Statistical analyses demonstrated that each site harbored phylogenetically distinct microbial populations that were more diverse at Ross site compare to winze site.

***Keywords*:** Homestake, DUSEL, gold mine, microbial diversity

The deep terrestrial subsurface environments such as those exemplified by ultra-deep mines represent an emerging area for exploring microbial populations with bewildering arrays of metabolic capabilities (DeFlaun *et al.*, 2007; Rastogi *et al.*, 2009). Culture-independent methods, in particular the recovery and analyses of 16S rRNA genes directly from environmental samples, provides a means of investigating microbial populations that eliminates the dependence on isolation of pure cultures (Pontes *et al.*, 2007). Culture-dependent and -independent surveys on gold mines in Japan (Inagaki *et al.*, 2003; Hirayama *et al.*, 2005; Nunoura *et al.*, 2005), South Africa (Takai *et al.*, 2001; Baker *et al.*, 2003; Onstott *et al.*, 2003; DeFlaun *et al.*, 2007), and the United States (Nemergut *et al.*, 2004; Rastogi *et al.*, 2009) have shown active microbial populations composed of diverse groups of microorganisms with unusual physiological properties. These reports also showed the existence of numerous novel and yet uncultured microbes representing a hitherto unidentified diversity. These mines harbor unique extreme environments for microorganisms, both natural and anthropogenic, including extreme temperature, pressure, low oxygen concentration, toxic metals, and pH.

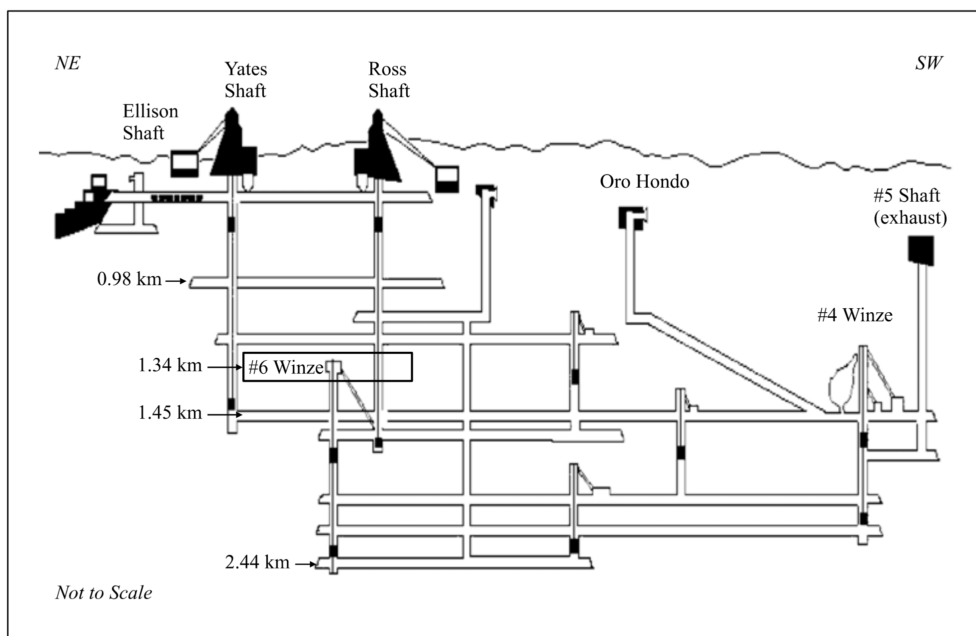
The Homestake gold mine (44°35′2074″N, 103°75′082″W) is the deepest mine (2.4 km deep) in the North America and had largest gold deposit ever found in the Western Hemisphere (Davis *et al.*, 2009; Rastogi *et al.*, 2009). On 10 July 2007, the National Science Foundation, USA has announced this mine as a site for Deep Underground Science and Engineering Laboratory (DUSEL). This former gold mine offers a unique opportunity for direct exploration of the mining-impacted deep subsurface environment. In a recent study, we have cultured the cellulose-degrading mesophilic and thermophilic bacteria from the Homestake gold mine (Rastogi *et al.*, 2009). However, to date culture-independent analysis of microbial diversity present in the Homestake mine have not been reported, and therefore the microbial communities existing in mining-impacted deep subsurface soil remains largely uncharacterized. Furthermore, only mineralogy data on the Homestake mine are available (Bachman and Caddey, 1990) but detailed geochemical characteristics of soils have not yet been reported. In the present study, our objective was to elucidate the bacterial and archaeal diversity in the soil samples of Homestake gold mine by constructing universal 16S rRNA gene libraries. The physical and chemical characteristics of soil samples were also studied in order to reveal biologically important chemical constituents. The results were compared with previous reports on microbial diversity assessments from gold mines and putative metabolic characteristics of retrieved phylotypes were inferred from the closest cultivable neighbors present in the database.

* To whom correspondence should be addressed.
(Tel) 1-605-394-1240; (Fax) 1-605-394-1232
(E-mail) Rajesh.Sani@sdsmt.edu

Materials and Methods

Subsurface soil collection

The location and geological setting of the Homestake mine have been described earlier (Bachman and Caddey, 1990; Davis *et al.*, 2009). The Homestake mine contains potentially acid-generating sulfide minerals approximately 8% by volume along with acid-neutralizing carbonate minerals including si-



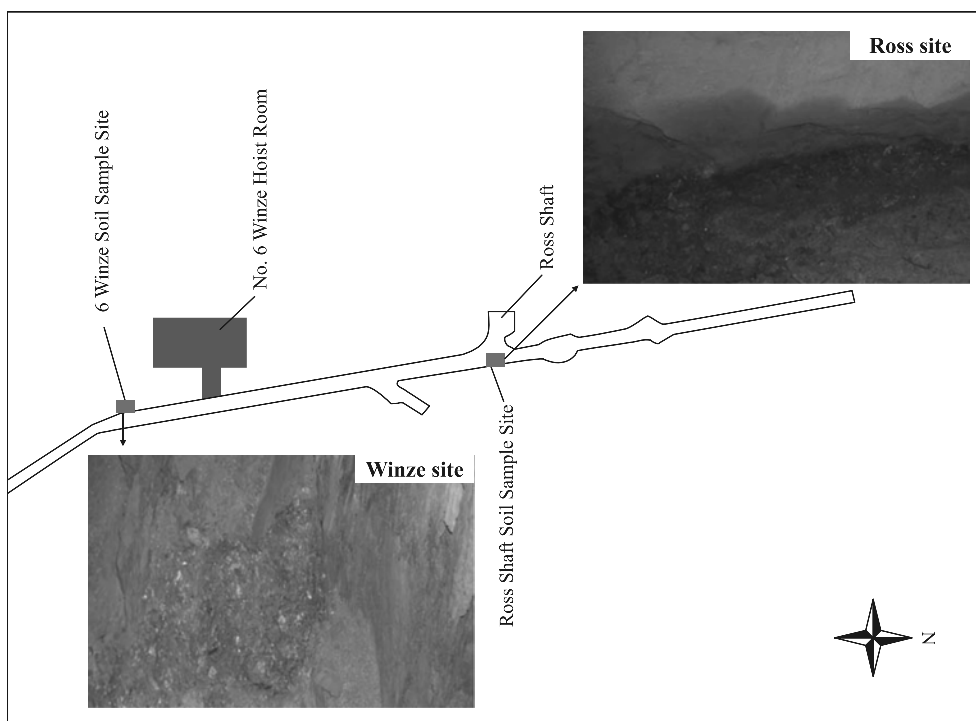


Fig. 1. (A) Schematic cross section and locations of the sampling sites (shown inside box) in the Homestake gold mine (Davis *et al.*, 2009). Soil samples were collected at 1.34 km depth. The boxed section has been shown in detail in Fig. 1B. (B) Detailed plane section of Ross and winze sites at 1.34 km depth. The area marked in Fig. 1A inside the box has been shown here in detail to show the actual sampling sites.

derite. The former Homestake mine operated continuously from 1876 to 2001 and contained 63 levels and about 515 km of underground workings, the deepest being 2.44 km beneath the surface (Fig. 1A). The Precambrian aquifer was pumped continuously during the operation of the mine and in June 2003, pumps were turned off. Current inflow of water in the underground working is about 2,839 L/min. Water enters the mine from the surface primarily through the shaft and airways that intersect with the open pit.

In May 2008, two soil samples were collected corresponding to the Ross shaft and No. 6 winze of the Homestake mine at a depth of 1.34 km when the water level was 0.02 km below this level (Fig. 1A). One sample was directly across the landing from the Ross shaft, one of two primary shafts from the surface into the mine, and one was outside the No. 6 winze hoist room (Fig. 1B). Both samples were collected along the junction of the drift wall and the floor, where a small accumulation of soil debris had built up through the years of mining. These areas were not disturbed by any type of activities including human trafficking from June 2003 to May 2008. The outer surfaces of the built up were discarded and only inner parts were collected using sterile spatulas for microbial diversity analyses. This was done to minimize the chances of contamination from exogenous microbes. Similar manual methods have been used earlier for the collection of water and rock samples to study the subsurface microbial diversity of gold mines in South Africa (Takai *et al.*, 2001) and Japan (Inagaki *et al.*, 2003; Hirayama *et al.*, 2005; Nunoura *et al.*, 2005). The temperature at the time of sampling was 26°C which was measured using a mercury thermometer. The samples were transported to laboratory in sterile polypropylene bottles on ice and stored at -20°C until analysis. Soil samples were homogenized in sterile pestle and motor inside a laminar flow hood, and then used for chemical characterization and DNA extraction.

Physical and chemical characterization of soil samples

Elemental compositions of soil samples were characterized at the Geo Analytical Laboratories (Washington State University, Pullman, WA, USA) using X-ray fluorescence spectroscopy (XRF) as described earlier (Rastogi *et al.*, 2008). Detailed physical and chemical characterizations of soils were determined at Soil and Plant Analysis Lab (University of Wisconsin, MD, USA). Since total metal concentration does not necessarily reflect the amount of metals that are bioavailable (water soluble), therefore water soluble concentration of selected metals were also determined. Major and minor elements, total minerals, and metals in soils were measured by ICP-OES (Inductively coupled plasma optical emission spectrometry) and ICP-MS (Inductively coupled plasma mass spectroscopy). All measurements were performed in duplicate and included appropriate controls. The detailed protocols for all analyses are described elsewhere (Peters, 2007).

DNA extraction, PCR, library construction, and phylogenetic analyses

Total DNA was extracted from 200 mg of each soil sample using a PowerSoil™ DNA Isolation kit (MO Bio, USA) according to the manufacturer's instructions. Polymerase chain reaction (PCR) amplification of the prokaryotic 16S rRNA genes was carried out with universal primers 530F and 1490R as described earlier (Rastogi *et al.*, 2009). A control PCR reaction was set up to check for any non-specific amplification. PCR products from triplicate reactions were pooled and concentrated for cloning reactions.

PCR products of expected size (~950 bp) were cloned in pGEM-T Easy Vector and transformed into *Escherichia coli* JM109 (Promega, USA) as per manufacturer's protocol. Plasmids were isolated using a Plasmid Extraction kit (QIAGEN, USA) and nucleotide sequences of cloned genes were determined in a 3730 DNA analyzer (Applied Biosystems, USA) for all the correct-sized clones in a library. The 16S rDNA sequences were checked for the presence of any anomalies by Mallard at 99.9% cut-off line (Ashelford *et al.*, 2006). The similarity search for the sequences were carried out by BLAST (N) program of the National Center of Biotechnology Information, MD, USA, and alignment was carried out by CLUSTAL W program available at European Molecular Biology Laboratory, Cambridge, UK. A Jukes-Cantor corrected distance matrix was calculated by DNADIST

Table 1. X-ray fluorescence spectroscopic analysis of major and trace elements of soil samples collected from the Ross and winze sites of the Homestake gold mine

Major elements (% weight)	Ross soil sample	Winze soil sample
SiO_2	38.06	52.69
TiO_2	0.252	0.412
Al_2O_3	5.96	8.16
FeO	22.38	6.03
MnO	0.385	0.078
MgO	2.97	5.18
CaO	5.96	6.70
Na_2O	0.35	0.45
K_2O	1.37	2.43
P_2O_5	0.084	0.127
Trace elements (mg/kg dry soil)		
Ni	60	73
Cr	119	99
Sc	7	9
V	57	165
Ba	253	459
Rb	44	85
Sr	139	254
Zr	54	85
Y	13	19
Nb	4.7	7.7
Ga	9	10
Cu	175	149
Zn	274	257
Pb	125	24
La	28	25
Ce	37	49
Th	7	6
Nd	17	22
U	3	6

Table 2. Physical and chemical characterization of Ross and winze soil samples

Physical/Chemical characteristics	Ross soil		Winze soil	
Organic matter (%)	1.3		0.7	
pH	6.6		6.7	
Sample density (g/cm^3)	1.26		1.08	
Soil texture	Sandy loam		Sandy loam	
Sand (%)	61		67	
Silt (%)	27		23	
Clay (%)	12		10	
Solids (%)	83.73		86.03	
Moisture (%)	16.27		13.97	
Soluble salts (salinity) (dS/m)	2.38		2.39	
Total carbon (%)	1.99		2.86	
Total sulfur (%)	2.66		1.79	
Total nitrogen (%)	0.06		0.06	
Ammonium-nitrogen (mg/kg)	16.07		<0.01	
Sulfate sulfur (mg/kg)	325.2		363.0	
Sodium (mg/kg)	131		107	
Phosphorus (mg/kg)	1		2	
Potassium (mg/kg)	90		69	
Calcium (mg/kg)	22283		2118	
Magnesium (mg/kg)	564		797	
Boron (mg/kg)	1.6		1.1	
Manganese (mg/kg)	31		50	
Zinc (mg/kg)	104.3		337.4	
Total minerals	(a)	(b)	(a)	(b)
P (%)	0.07	<0.001	0.02	<0.001
K (%)	0.54	0.01	0.42	0.01
Ca(%)	2.85	0.38	2.52	0.56
Mg (%)	1.28	0.11	1.87	0.05
S (%)	2.04	0.39	1.49	0.57
Zn (mg/kg)	307.39	0.08	1040.19	<0.01
B (mg/kg)	3.92	0.31	5.51	0.43
Mn (mg/kg)	1993.61	0.45	704.27	0.16
Fe (mg/kg)	75657.3	1.4	51539.7	0.1
Cu (mg/kg)	81.55	0.16	96.11	0.11
Al (mg/kg)	16704.2	<0.5	10222.7	0.7
Na (mg/kg)	252.7	213.1	238.3	92.1
Heavy metals (mg/L)	(a)	(b)	(a)	(b)
Cd	9.60	<0.04	3.37	<0.04
Co	9.80	<0.03	14.39	<0.03
Cr	74.331	<0.01	39.771	<0.01
Cu	81.55	0.16	96.11	0.11
Fe	75657	1.378	51540	0.103
Mn	1993.6	0.449	704.3	0.156
Mo	4.00	<0.04	4.25	0.07
Ni	<0.3	<0.03	<0.03	<0.03
Pb	93.76	<0.2	29.17	<0.2
Zn	307.39	0.08	1040	<0.01
Li	20.520	0.181	21.422	0.275
As	670.8	<0.3	177.5	<0.3
Se	<3	<0.3	<3	<0.3
Water soluble anions (mg/L)				
F^-	0.03		<0.01	
Cl^-	21.9		17.2	
Br^-	<0.01		<0.01	
NO_3^-	39.1		15.7	
PO_4^-	<0.02		<0.02	
SO_4^-	15156		9237	

(a) Concentration determined after digestion with concentrated nitric acid
(b) Concentration determined after extraction with water

program of PHYLIP (Felsenstein, 1989). Distance matrix was used to assign sequences in various operational taxonomic units (OTU) or phylotypes using DOTUR at 97% sequence similarity cut-off (Schloss and Handelsman, 2005). The phylogenetic trees were constructed by the neighbor-joining method (with 1,000 bootstrap replicates) using MEGA v 3.1 (Kumar *et al*., 1993). Phylogenetic affiliations of OTUs were confirmed using naïve Bayesian classifier available at the Ribosomal Database Project site (Wang *et al*., 2007).

Statistical analyses of microbial diversity

Good's coverage (C) of library was calculated as [1-(n/N)] ×100, where 'n' is the number of single clone OTUs and 'N' is the library size (Schloss and Handelsman, 2005). Statistical parameters including randomized rarefaction curves and Shannon-Weaver (*H*') index, were estimated to quantify the diversity of phylotypes using DOTUR. Rarefaction curves were plotted to assess the library coverage at phylum, class, order, genus, and species levels at sequence similarity cut-off values of 80, 85, 90, 95, and 97%, respectively using DOTUR. The fraction of OTUs shared between two sites was estimated using SONS program at 97% sequence similarity cut-off (Schloss and Handelsman, 2006a). LIBSHUFF v 0.96 was used to compare the libraries according to the specifications given at the LIBSHUFF website (Henriksen, 2004). Two libraries were considered significantly different in community composition at critical *P*-value ≤ 0.025 with a confidence of 95% ($P=0.05$).

Nucleotide sequence accession numbers

The 16S sequences belonging to the Ross and winze libraries are available in the GenBank under the accession numbers FJ184605-FJ184769 and FJ184770-FJ184934, respectively.

Results

Geochemistry of soil samples

The physical and chemical characteristics of soil samples collected from the Ross and winze sites are presented in Table 1 and 2. Significant differences were observed in the metal composition and physical properties of Ross and winze soil samples. Both soil samples contained significant amount of toxic metals such as As, Cd, Co, Cr, Cu, Ni, Pb, and Zn. However, the water soluble concentrations of these toxic metals were very low. Since the gold deposits at the Homestake mine are typically associated with the iron formation (Bachman and Caddey, 1990) therefore both soil samples contained very high amount of iron (Table 2). Presence of different water soluble ions (sulfate, nitrate, iron) in soils indicates that these can support the growth of various chemotrophic microorganisms in deep subsurface environment of the Homestake mine where the energy sources are limited. Lack of the organic energy sources in the Homestake mine may particularly select the chemolithoautotrophs with novel metabolic capabilities.

Phylogenetic affiliation of OTUs retrieved from Ross site

Phylogenetic analysis of 165 clone sequences generated 110 OTUs that were affiliated to *Bacteria* (107 OTUs) and *Archaea* (3 OTUs) domains. The phylogenetic affiliation of OTUs among different taxonomic groups is indicated in Fig. 2A and B. Majority of the OTUs (>97%) identified at Ross site were related to uncultured bacterium clones recovered from a variety of environments including acid mine biofilms, mine-tailings, metal-contaminated soils, and subsurface soils and waters. A total of 23 clones (14% of total clone diversity) were clustered with sequences belonging to class *Thermoprotei* in *Crenarchaeota* and showed similarities with uncultured archaeal clones recovered from rhizosphere (EU309860) and acid mine biofilm (AY082453) (Fig. 2A). Phylotypes within the *Bacteria* were clustered in phyla, namely *Proteobacteria* (81 clones), *Acidobacteria* (22 clones), *Planctomycetes* (11 clones), *Actinobacteria* (4 clones), *Firmicutes* (4 clones), *Bacteroidetes* (4 clones), *Chloroflexi* (4 clones), *Gemmatimonadetes* (2 clones), *Nitrospirae* (2 clones), *Verrucomicrobia* (1 clone), *Chlorobi* (1 clone), and candidate divisions TM7 (1 clone) and OP10 (1 clone) (Fig. 2A and B). A total of 4 clones represented by OTU-7, 36, 48, and 54 were considered as unclassified because they formed distinct clusters due to low similarities with database sequences. A single clone represented by OTU-28 within *Firmicutes* was similar to *Paenibacillus* sp. (AM162342). *Proteobacteria* was the most predominant group in the library (49% of total clone diversity). Within the *Proteobacteria*, all clones were distributed in four classes: α-(18 clones), β-(7 clones), γ-(44 clones), and δ-*Proteobacteria* (4 clones) (Fig. 2B). A total of 8 clones within the *Proteobacteria* were considered as unclassified because of their low similarity with known proteobacterial lineages. Majority of the *Proteobacteria*-affiliated clones (>98%) were similar to uncultured bacterium clones; however, phylotypes represented by OTU-81 and 82 formed sister lineages with cultivable proteobacterial lineages namely *Zobellella* (DQ195676) and *Rheinheimera* (DQ985066), respectively.

Phylogenetic affiliation of OTUs retrieved from winze site

Phylogenetic analysis of 165 clone sequences yielded 100 OTUs that were clustered in the *Bacteria* (96 OTUs) and *Archaea* (4 OTUs) domains. The phylogenetic position and distribution of OTUs are indicated in Fig. 3A and B. A total of 10 clones (6% of total clone diversity) represented by 4 OTUs belonged to class *Thermoprotei* within *Crenarchaeota*. These clones were phylogenetically related to uncultured archaeal clones recovered from water samples of the South African and Japanese gold mines and Altamira cave in Spain (Fig. 3A). Within the *Bacteria*, OTUs were grouped into 7 phyla represented by *Proteobacteria* (81 clones), *Actinobacteria* (26 clones), *Acidobacteria* (5 clones), *Planctomycetes* (12 clones), *Firmicutes* (6 clones), *Bacteroidetes* (3 clones), *Verrucomicrobia* (2 clones), and candidate division TM7 (2 clones). These phyla were predominantly composed of uncultured bacterium clone sequences reported from soils, sediments, and mine-tailings. A total of 13 OTUs containing 18 clones were not related to any lineage with known taxonomic affiliation and were considered as unclassified. The *Firmicutes* group contained a lineage related with *Alkalibacillus silvisoli* (AB264528). Among 81 clones related to *Proteobacteria*, 23

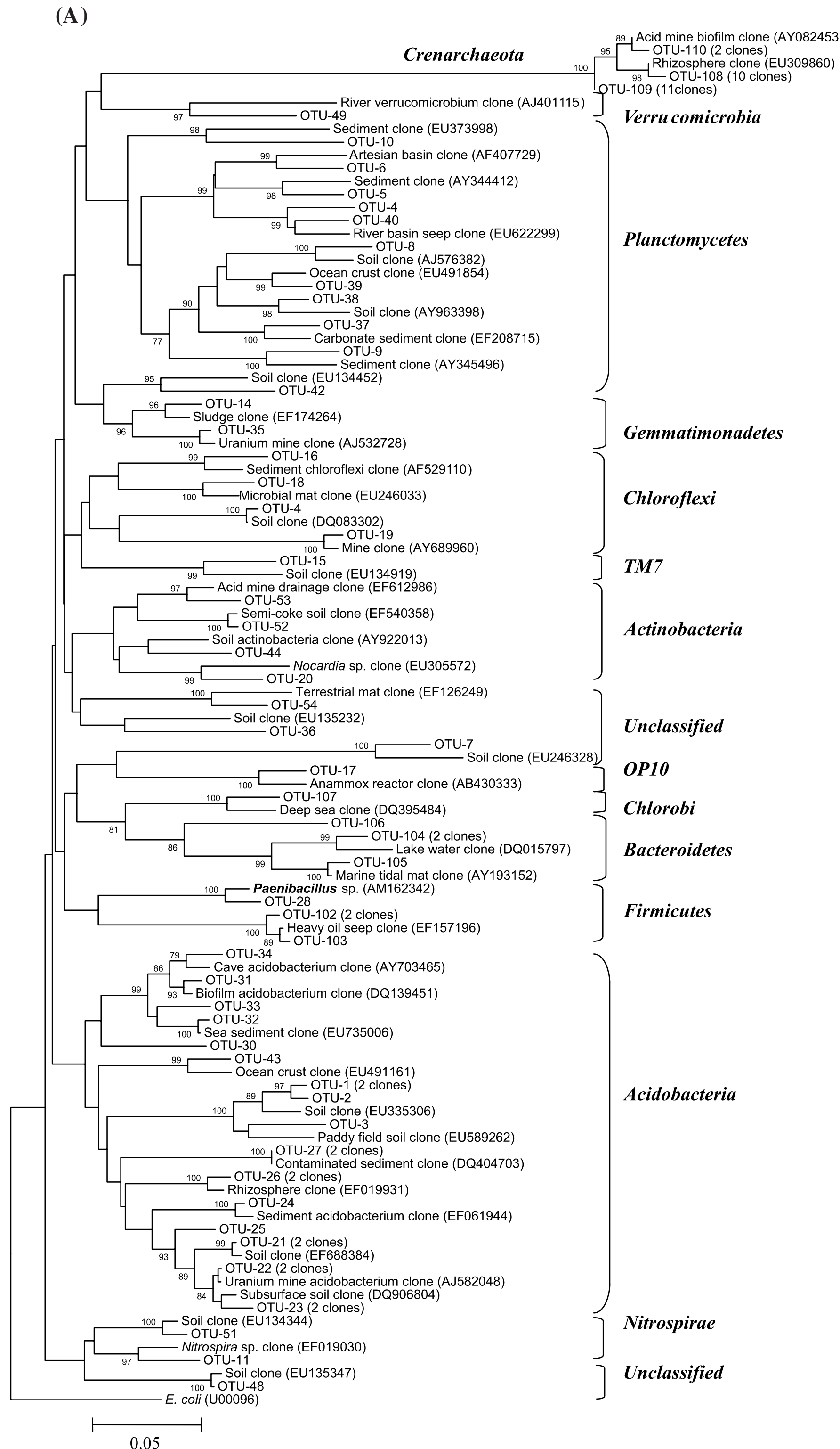


Fig. 2. (A) Phylogenetic dendrogram of 16S rRNA gene sequences (excluding *Proteobacteria*) derived from the Ross site with reference sequences in GenBank. For OTUs representing multiple clones, the number of additional clones is given in parentheses. Phylotypes that were most closely related to cultivable lineages are indicated by bold letters. *E. coli* (U00096) was selected as out-group to root the tree. The *scale bar*represents 0.05 substitutions per nucleotide position. Bootstrap values greater than 75% are indicated at the nodes. (B) Phylogenetic dendrogram of *Proteobacteria*-related 16S rRNA gene sequences derived from the Ross site with reference sequences in GenBank. Phylotypes that were most closely related to cultivable lineages are indicated by bold letters. *Sulfolobus acidocaldarius* (D14053) was selected as out-group to root the tree. The tree was constructed as described in the legend to Fig. 2A.

Fig. 2. Continued

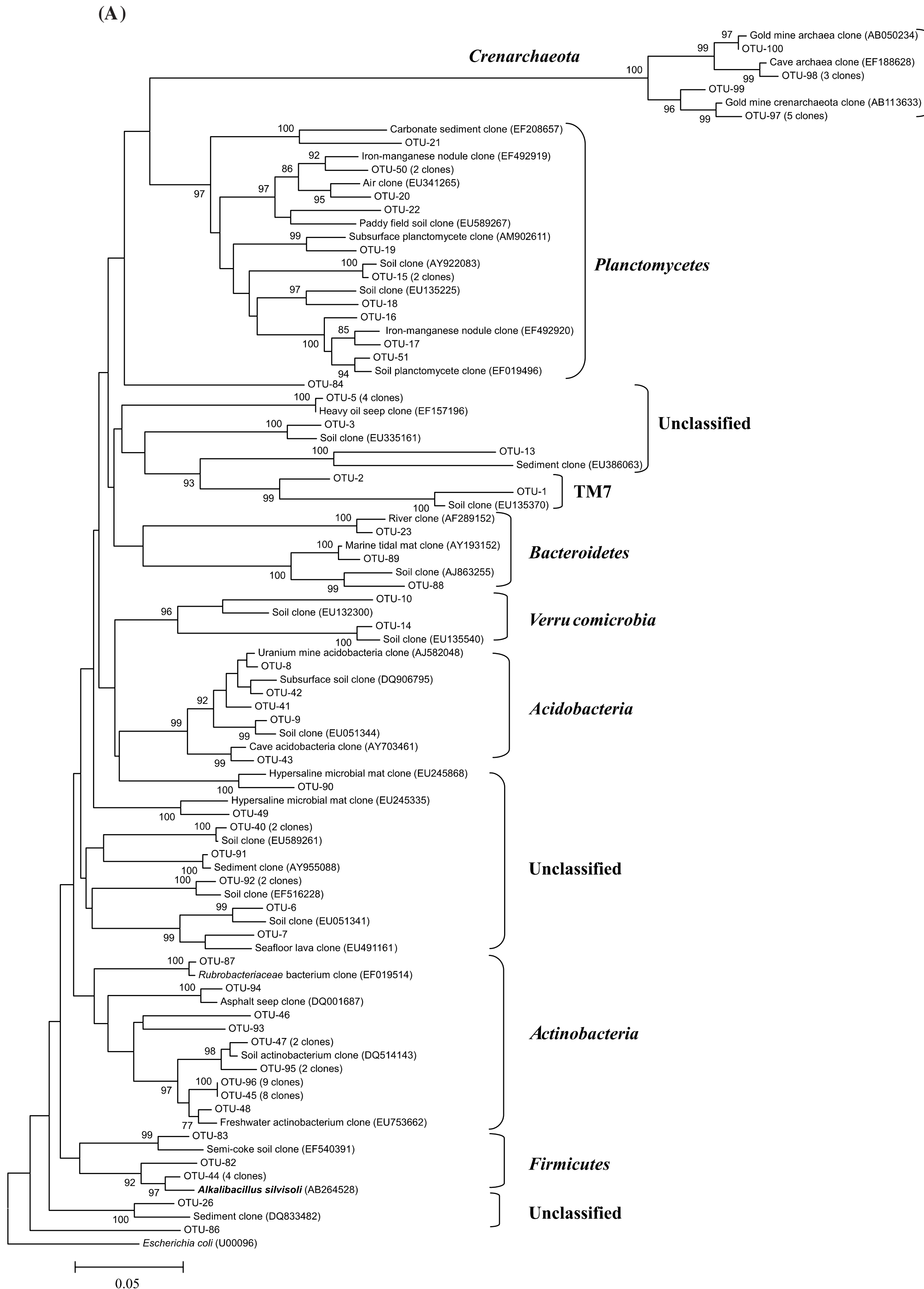


Fig. 3. (A) Phylogenetic dendrogram of 16S rRNA gene sequences (excluding *Proteobacteria*) derived from the winze site with reference sequences in GenBank. Phylotypes that were most closely related to cultivable lineages are indicated by bold letters. The tree was constructed as described in the legend to Fig. 2A. (B) Phylogenetic dendrogram of *Proteobacteria*-related 16S rRNA gene sequences derived from the winze site with reference sequences in GenBank. Phylotypes that were most closely related to cultivable lineages are indicated by bold letters. The tree was constructed as described in the legend to Fig. 2A.

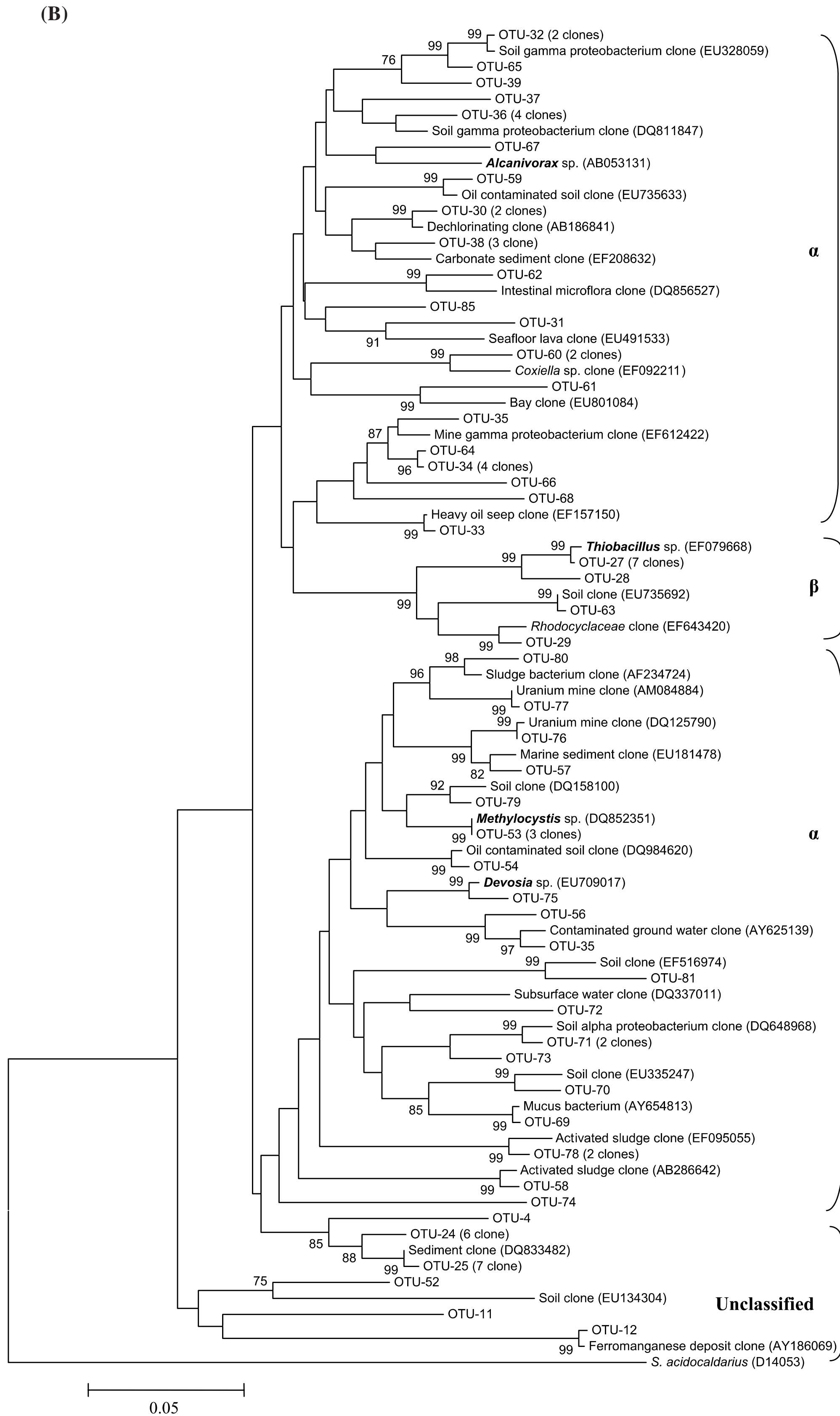


Fig. 3. Continued

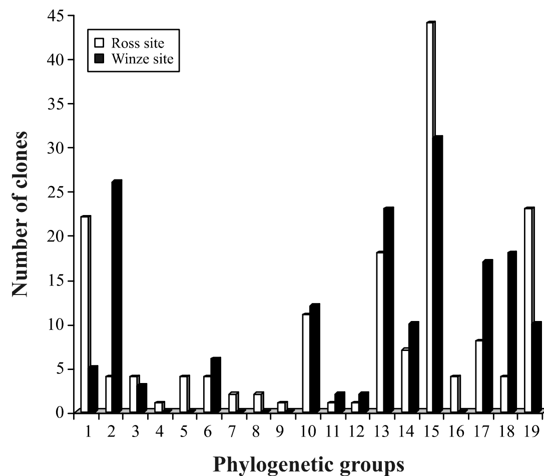


Fig. 4. Bar diagram showing the relative distribution of 16S rRNA gene clones in the Ross and winze sites. The clones were assigned to a particular group based on the phylogenetic analyses. Columns: 1, *Acidobacteria*; 2, *Actinobacteria*; 3, *Bacteroidetes*; 4, *Chlorobi*; 5, *Chloroflexi*; 6, *Firmicutes*; 7, *Gemmatimonadetes*; 8, *Nitrospirae*; 9, Candidate division OP10; 10, *Planctomycetes*; 11, Candidate division TM7; 12, *Verrucomicrobia*; 13, α-*Proteobacteria*; 14, β-*Proteobacteria*; 15, γ-*Proteobacteria*; 16, δ-*Proteobacteria*; 17, Unclassified *Proteobacteria*; 18, Unclassified clones; 19, *Crenarchaeota.*

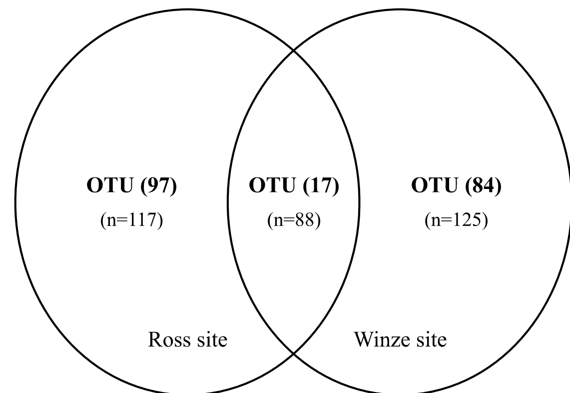


Fig. 5. Venn diagram showing fraction of similar OTUs observed between Ross and winze sites. The numbers in the circles indicate the estimated numbers of OTUs (and the number of clone sequence) that are shared or unique to each site. The Venn diagram was plotted using SONS program as described in text.

clones were affiliated to α-, 10 clones to β-, and 31 clones to γ-*Proteobacteria* (Fig. 3B). A total of 17 clones spanning in 6 OTUs were not affiliated to any known proteobacterial lineages and hence considered as unclassified *Proteobacteria*. Within the α-*Proteobacteria* cluster, two sequences represented by OTU-75 and OTU-53 were phylogenetically similar to *Devosia* sp. (EU709017) and *Methylocystis* sp. (DQ852351), respectively. Within the β-*Proteobacteria*, OTU-27 was recovered that showed related phylogeny with *Thiobacillus* sp. (EF079668). A detected γ-*Proteobacteria*-related clone (OTU-67) showed similarity with *Alcanivorax* sp. (AB053131; a petroleum-degrading marine bacterium).

Statistical analysis of 16S rRNA gene libraries

Clone library comparison by LIBSHUFF yielded a critical p-value of 0.001 which confirmed significant differences in the microbial community composition between the Ross and winze sites. Figure 4 shows the relative distribution of different taxonomic groups retrieved from Ross and winze sites. SONS analysis quantified the degree of overlap between OTU memberships and structures of two sites. The program estimated a total of 198 OTUs from the combined dataset (330 clone sequences). Out of total 198 OTUs detected, 17 OTUs (containing 88 sequences) were common at both the sites. Thus despite of differences in the physical and chemical composition, both sites had shared phylotypes diversity (Fig. 5). DOTUR analysis showed that sample collected from the Ross site (110 OTUs) had relatively more species richness than the winze site (100 OTUs) as indicated by the number of phylotypes retrieved from the same library sizes (N=165). Shannon diversity index pointed relatively higher species diversity at the Ross site (H'=4.43) compared to the winze site (H'=4.32).

Rarefaction analyses of libraries generated non-asymptotic plots (θ>0°) for each taxonomic level (Fig. 6A and B). However, rarefaction curves at phylum and class levels showed that microbial diversity was sampled with greater coverage at these two taxonomic levels. This was also evident from corresponding Good's coverage values which were relatively higher at phylum and class levels for both libraries. Microbial diversity at other taxonomic levels (order, genus, and species) was not sampled exhaustively with current sequencing efforts. The lower diversity coverage was due to the fact that both libraries contained primarily single clone OTUs (rare phylotypes). The greater the number of these single clone OTUs in a library the lesser will be the coverage leading to a more non-asymptotic curve. The non-asymptotic curves as obtained in this study is not an unusual observation because it is difficult to get a complete species census [asymptotic rarefaction curves (θ=0°)] in an environmental sample and there are always undiscovered species left in almost every molecular inventory (Schloss and Handelsman, 2006b; DeSantis *et al*., 2007). Literature has also shown that in an environmental sample precise estimation of species composition will require >10^4 clone sequences per sample (DeSantis *et al*., 2007). Thus most of the published reports utilizing cloning and sequencing approaches represent only a partial coverage of the microbial diversity present in a sample. Nevertheless, clone libraries of 16S rRNA genes definitely permit an initial survey of the microbial diversity in any given sample. Clone libraries containing less number of single-clone OTUs yield more strongly curvilinear rarefaction plots. Therefore the clone library of winze site (n=76) yielded more strongly curvilinear plots than the Ross site (n=85) at each taxonomic level (Fig. 6A and B). Rarefaction analyses were in strong correlation with Good's coverage and confirmed that at each taxonomic level, winze clone library covered more diversity than Ross clone library.

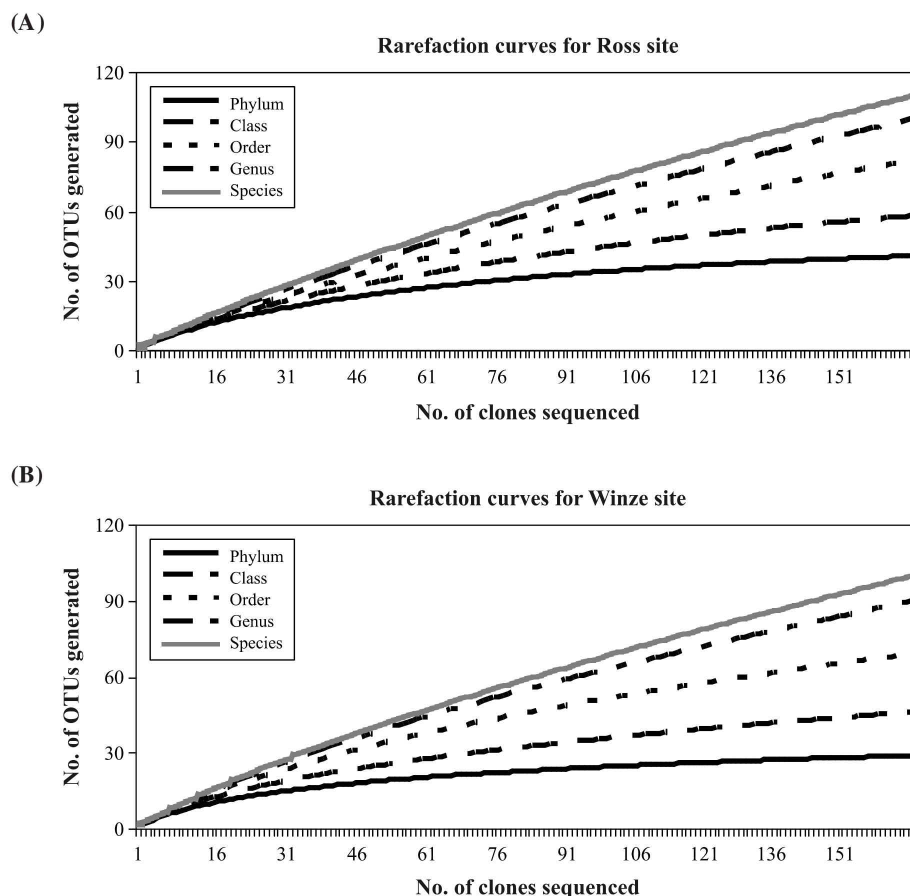


Fig. 6. (A) Rarefaction curves of 16S clone library of Ross site. The curves illustrate the relationship between number of clones sequenced and the number of OTUs observed at different taxonomic levels. The steepness (θ) of each curve was calculated from equation y=mx+c where 'x' and 'y' are the coordinates of the points 'm' is the gradient and 'c' is the 'y intercept' of the straight line graph. In a molecular inventory where 100% diversity is captured θ reaches to 0 i.e., the curve takes the form of plateau. The 'θ' values for phylum, class, order, genus, and species level were 12, 18, 25, 30, and 33°, respectively. Estimated Good's coverage (C) for phylum class, order, genus, and species level were 93, 84, 65, 54, and 49%, respectively. (B) Rarefaction curves of 16S clone library of winze site. The curves were plotted as described in the legend to Fig. 6A. The 'θ' values for phylum, class, order, genus, and species level were 8, 14, 21, 27, and 30°, respectively. Estimated coverage for phylum class, order, genus, and species level were 95, 89, 77, 62, and 54%, respectively.

Discussion

Microbial community composition in metal-contaminated soils

Previous microbial diversity studies on gold mines demonstrated *Proteobacteria* to constitute a major proportion of the clone libraries (Onstott *et al*., 2003; Nemergut *et al*., 2004). Our results corroborated with the previous reports as lineages belonging to *Proteobacteria* constituted the most abundant group at both the sites. Species belonging to *Proteobacteria* are well-known to survive in low-nutrient environments, metal reduction, and metal resistance (Akob *et al*., 2007). Comparing our findings on microbial diversity with other gold mines (Takai *et al*., 2001; Inagaki *et al*., 2003; Onstott *et al*., 2003; Nemergut *et al*., 2004; Hirayama *et al*., 2005; Nunoura *et al*., 2005), our results showed remarkable similarities at phylum level. The bacterial phyla in common were *Acidobacteria*, *Actinobacteria*, *Bacteroidetes*, *Chlorobi*, *Chloroflexi*, *Firmicutes*, *Gemmatimonadetes*, *Nitrospirae*, *Proteobacteria*, and candidate divisions OP10 and TM7. However, differences were observed in the analysis of specific lineages within these phyla, such as the presence of genera e.g., *Rheinheimera*, *Zobellella*, and *Devosia*, in the Homestake mine that have not been documented in earlier studies from any gold mines. Furthermore, previous studies on microbial diversity from gold mines did not document *Verrucomicrobia* as found in our study. These overall differences suggest that microbial communities in gold mines are site specific due to unique geological and physicochemical settings and diverse microbial communities were adapted to elevated levels of metal-contamination after long stress. However, several phylotypes at Ross and winze sites showed strong relatedness with clones previously reported from gold or uranium mines and metal-contaminated sites. In addition SONS analysis showed that there was similar phylotype diversity between Ross and winze sites. The Homestake gold mine contained a high amount of toxic metals including U; however, this mine is geographically distinct from previously studied gold

or uranium mines or metal-contaminated sites. Thus it was interesting to note such similarities in retrieved phylotypes of the Homestake gold mine with other gold or uranium mines or metal-contaminated sites. This may suggest that such common lineages are probably performing similar processes as they are at the other gold mines, and may represent important indigenous bacteria of gold mines or metal-contaminated sites.

Inference of putative metabolic characteristics of OTUs

In the present study several phylotypes could not be classified to any known taxonomic group suggesting that they might represent new taxa and their role in mine ecology remains unclear. Though the phylogenetic affiliation of phylotypes displaying low sequence similarities to any sequence in the database can be estimated from a phylogenetic tree, the extrapolation of their metabolic characteristics from well-characterized cultured species to distantly related phylotypes may not always be reasonable. Another problem in inferring metabolic characteristics arises when phylotypes were related to phyla for which no cultivated representatives are known. For example at the Ross and winze sites, phylotypes similar to sequences from candidate divisions OP10 and TM7 were retrieved. To date no cultivable lineages belonging to these candidate divisions have been isolated in cultures. These bacterial divisions are exclusively represented by environmental sequence data and are unstudied (Janssen, 2006). Other phyla such as *Acidobacteria* and *Verrucomicrobia* contain only few cultured members (Janssen, 2006). Therefore their physiological roles in the natural environments including in mine ecology remain unknown. Any physiological interpretations must therefore rely heavily on the environmental settings of a sample from which a phylotype is isolated. Our study has avoided over interpretation of data based solely on clone sequences. The putative metabolic characteristics of only those phylotypes were elucidated that showed strong phylogeny supported by high bootstrap values (>50%) and BLAST identity (>90%) to some cultivable lineage in database.

For example, at Ross site a lineage showed 100% bootstrap value and 95% BLAST identity with *Paenibacillus* sp. Members of the *Paenibacillus* are facultative anaerobes, and have been shown to produce metal-binding exopolysaccharides that can trap Cd, Cu, Co, Ni, Pb, and Zn (Morillo *et al*., 2006). Genus *Paenibacillus* has been widely reported to utilize cellulose as a sole source of carbon (Wang *et al*., 2008). During active mining-operations for over 125 years, surface (exogenous) microbes and lignocellulosic substrates were introduced into the extreme deep subsurface environment of the Homestake mine. In another study, we have isolated cellulose-degrading mesophilic *Paenibacillus* strains from deep subsurface soil collected from the Homestake mine (Rastogi *et al*., 2009). Some species of *Paenibacillus* have also been shown in literature to possess nitrogen fixing capabilities (Rosado *et al*., 1998). Another phylotype at Ross site showed strong grouping with *Rheinheimera* sp. (identified from marine environment) supported by 99% BLAST identity and 100% bootstrap value. This genus has not been identified from mine or metal-contaminated environments and therefore its putative role in mine ecology is unclear. Additionally at Ross site, a phylotype showed 97% BLAST identity and a node bootstrap value of 100% with *Zobellella taiwanensis* (DQ195676) a heterotrophic, fermentative facultative anaerobe isolated from sediments. This bacterium has been shown to reduce nitrate or nitrite to the gaseous end product(s) nitrous oxide (N_2O) or N_2 (Lin and Shieh, 2006). Presence of water soluble nitrate ions (39.1 mg/L) as detected at Ross site may act as electron acceptors for the growth these bacteria.

At winze site, a phylotype showed similarity (93% sequence identity, 99% bootstrap support) with *Thiobacillus* sp. that has been isolated from freshwater sediment. In general, *Thiobacillus* have been recovered from gold mine tailings (Blowes *et al*., 1998). *Thiobacilli* are chemoautolithotrophic, sulfur-oxidizing bacteria that are restricted to habitats where both an electron donor [reduced sulfur compounds (S, H_2S, and $S_2O_3^{-2}$) or Fe(II) in some cases] and an electron acceptor (O_2 or NO_x) simultaneously exists. These bacteria produce sulfuric acid and Fe(III) as by-products of their metabolism and play very important role in biomineralization (Southam and Beveridge, 1992). Southam and Beveridge (1992) further demonstrated a very strong association between *Thiobacillus* species and the sulfide minerals, which helps account for their prominence in tailings environments. Sulfides and pyrites (FeS_2) are commonly associated with gold deposits which can serve as electron donors for the growth of these bacteria in deep subsurface of the Homestake mine (Bachman and Caddy, 1990). In addition, chemical characterization of winze soil also showed significant amount of sulfur (1.79%) which can act as energy source for these bacteria. Another phylotype at winze site was similar to *Alcanivorax* sp. and exhibited 93% identity supported by 57% bootstrap node value. Member belonging to *Alcanivorax* are aerobic bacteria that exclusively uses aliphatic hydrocarbons as the sole source of carbon and energy (Liu and Shao, 2005). This genus is also known to grow anaerobically using nitrate as electron acceptor. Winze soil sample contained 15.7 mg/L water soluble nitrate ions which can act as electron acceptor. Phylotype represented by OTU-75 formed a common group supported by 99% bootstrap node value and displayed 95% similarity to *Devosia* sp. recovered from marine water. Genus *Devosia* have been reported to include members with nitrogen-fixing capabilities (Vanpyres *et al*., 2005). Sequences related to methane-oxidizing bacteria of genus *Methylocystis* (99% BLAST identity, 99% bootstrap value) were identified that grow on methane as their sole source of carbon and energy. *Methylocystis*-related phylotypes have been previously identified from Japanese gold mines (Hirayama *et al*., 2005). One of the retrieved lineages was 98% similar (94% bootstrap value) to *Alkalibacillus silvisoli*, an alkaliphilic (pH 7~10) bacteria isolated from forest soil.

In summary, the present study showed diverse microbial populations including yet-uncultivated microbial communities in the mining-impacted deep subsurface environment of the Homestake gold mine. Microbial diversity was found to be site specific and phylogenetically distinct. Molecular oxygen and surface microbes were introduced into the Homestake mine by human activities during mining-operations, thus the microbial community in soil samples does not nec-

essarily reflects an indigenous subsurface microbial population. However, various environmental constraints in the deep subsurface of mine would shape the microbial diversity as a novel and unique subsurface microbial ecosystem. The results on microbial diversity and geochemistry will serve as a vital comparison for future assessment of changes in microbial diversity and geochemistry as re-entry in Homestake mine continues and the deeper levels become exposed during the construction of the DUSEL.

PCR-based molecular ecological surveys have been used widely to investigate microbial diversity in variety of environments including gold mines. Nevertheless, considering the probability of various PCR-induced anomalies, differences in 16S rRNA gene copy numbers, methodological biases, and small clone library sizes (Pontes *et al*., 2007), it may be erroneous to assume that the obtained phylotype distribution would be the actual species distribution in subsurface soils. It should also be noted that 16S rRNA gene based phylogenetic analysis could greatly underestimate the microbial diversity. However, it provides information on some of the more significant species present in a given environment. In addition, the community analysis using molecular methods alone is not sufficient for predicting the metabolic functions within the environment.

In our study, especially the facts such as less than 3% of total clones (330) were related to known cultivable species and similarities of clones to phyla with no or few cultivable representatives further complicated this situation. Nevertheless, a molecular inventory of species is the first step to describe the unique microbial communities and forms basis for the development of improved culturing methods, and subsequently, understanding their metabolic roles. Research in our laboratory is currently underway where we are using high-density universal 16S microarrays for a comprehensive view of bacterial and archaeal community compositions existing in the Homestake gold mine.

Acknowledgements

This research was funded by the SDBOR Competitive Research Grant Award No. SDBOR/SDSMT 2010-09-05 for the project title “Generating Preliminary Microbial Data on Homestake Gold Mine”. The authors are grateful to Dr. Parag Vaishampayan (Jet Propulsion Laboratory, Pasadena, CA) for his help in doing phylogenetic analyses. We also would like to thank the anonymous reviewers whose critiques were instrumental in making our manuscript of an excellent quality.

References

Akob, D.M., H.J. Mills, and J.E. Kostka. 2007. Metabolically active microbial communities in uranium-contaminated subsurface sediments. *FEMS Microbiol. Ecol*. 59, 95-107.

Ashelford, K.E., N.A. Chuzhanova, J.C. Fry, A.J. Jones, and A.J. Weightman. 2006. New screening software shows that most recent large 16S rRNA gene clone libraries contain chimeras. *Appl. Environ. Microbiol.* 72, 5734-5741.

Bachman, R.L. and S.W. Caddey. 1990. The Homestake Mine, Lead, South Dakota: an overview, vol. 7, p. 89-94. *In* C.J. Paterson and A.L. Lisenbee (eds.), Metallogeny of Gold in the Black Hills, South Dakota. Society of Economic Geologists, Guidebook Series.

Baker, B.J., D.P. Moser, B.J. MacGregor, S. Fishbain, M. Wagner, N.K. Fry, B. Jackson, N. Speolstra, S. Loos, K. Takai, B.S. Lollar, J. Fredrickson, D. Balkwill, T.C. Onstott, C.F. Wimpee, and D.A. Stahl. 2003. Related assemblages of sulphate-reducing bacteria associated with ultradeep gold mines of South Africa and deep basalt aquifers of Washington State. *Environ. Microbiol*. 5, 267-277.

Blowes, D.W., J.L. Jambor, C.J. Hanton-Fong, L. Lortie, and W.D. Gould. 1998. Geochemical, mineralogical and microbiological characterization of a sulphide-bearing carbonate-rich gold-mine tailings impoundment, Joutel, Quebec. *Appl. Geochem.* 13, 687-705.

Davis, A., W. Roggenthen, L. Stetler, Z. Hladysz, and C. Johnson. 2009. Post-closure flooding of the Homestake mine at Lead, South Dakota. *Mining Engineering* March p. 43-47.

DeFlaun, M.F., J.K. Fredrickson, H. Dong, S.M. Pfiffner, T.C. Onstott, D.L. Balkwill, S.H. Streger, E. Stackebrandt, S. Knoessen, and E. van Heerden. 2007. Isolation and characterization of a *Geobacillus thermoleovorans* strain from an ultra-deep South African gold mine. *Syst. Appl. Microbiol.* 30, 152-164.

DeSantis, T.Z., E.L. Brodie, J.P. Moberg, I.X. Zubieta, Y.M. Piceno, and G.L. Andersen. 2007. High-density universal 16S rRNA microarray analysis reveals broader diversity than typical clone library when sampling the environment. *Microb. Ecol.* 53, 371-383.

Felsenstein, J. 1989. PHYLIP-Phylogeny inference package (version 3.2). *Cladistics* 5, 164-166.

Henriksen, J.R. 2004. webLIBSHUFF (http://libshuff.mib.uga.edu)

Hirayama, H., K. Takai, F. Inagaki, Y. Yamato, M. Suzuki, K.H. Nealson, and K. Horikoshi. 2005. Bacterial community shift along a subsurface geothermal water stream in a Japanese gold mine. *Extremophiles* 9, 169-184.

Inagaki, F., K. Takai, H. Hirayama, Y. Yamato, K.H. Nealson, and K. Horikoshi. 2003. Distribution and phylogenetic diversity of the subsurface microbial community in a Japanese epithermal gold mine. *Extremophiles* 7, 307-317.

Janssen, P.H. 2006. Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Appl. Environ. Microbiol.* 72, 1719-1728.

Kumar, S., K. Tamura, and M. Nei. 1993. MEGA: Molecular evolutionary genetics analysis. Pennsylvania State University, University Park, PA, USA.

Lin, Y.T. and W.Y. Shieh. 2006. *Zobellella denitrificans* gen. nov., sp. nov. and *Zobellella taiwanensis* sp. nov., denitrifying bacteria capable of fermentative metabolism. *Int. J. Syst. Evol. Microbiol.* 56, 1209-1215.

Liu, C. and Z. Shao. 2005. *Alcanivorax dieselolei* sp. nov., a novel alkane-degrading bacterium isolated from sea water and deep-sea sediment. *Int. J. Syst. Evol. Microbiol.* 55, 1181-1186.

Morillo, J.A., M. Aguilera, A. Ramos-Cormenzana, and M. Monteoliva-Sánchez. 2006. Production of a metal-binding exopolysaccharide by *Paenibacillus jamilae* using two-phase olive-mill waste as fermentation substrate. *Curr. Microbiol.* 53, 189-193.

Nemergut, D.R., A.P. Martin, and S.K. Schmidt. 2004. Integron diversity in heavy-metal-contaminated mine tailings and inferences about integron evolution. *Appl. Environ. Microbiol.* 70, 1160-1168.

Nunoura, T., H. Hirayama, H. Takami, H. Oida, S. Nishi, S. Shimamura, Y. Suzuki, F. Inagaki, K. Takai, K.H. Nealson, and K. Horikoshi. 2005. Genetic and functional properties of uncultivated thermophilic crenarchaeotes from a subsurface gold mine as revealed by analysis of genome fragments. *Environ. Microbiol*. 7, 1967-1984.

Onstott, T.C., D.P. Moser, S.M. Pfiffner, J.K. Fredrickson, F.J.

Brockman, T.J. Phelps, D.C. White, A. Peacock, D. Balkwill, R. Hoover, L.R. Krumholz, M. Borscik, T.L. Kieft, and R. Wilson. 2003. Indigenous and contaminant microbes in ultradeep mines. *Environ. Microbiol.* 5, 1168-1191.

Peters, J. 2007. Wisconsin Procedures for Soil Testing, Plant Analysis and Feed & Forage Analysis, No. 6, Soil Fertility Series. Soil Science Department, College of Agriculture and Life Sciences, University of Wisconsin (MD, USA). Available online at http://uwlab.soils.wisc.edu/procedures.htm

Pontes, D.S., C.I. Lima-Bittencourt, E. Chartone-Souza, and A.M. Amaral Nascimento. 2007. Molecular approaches: advantages and artifacts in assessing bacterial diversity. *J. Ind. Microbiol. Biotechnol.* 34, 463-473.

Rastogi, G., G.L. Muppidi, R.N. Gurram, A. Adhikari, K.M. Bischoff, S.R. Hughes, W.A. Apel, S.S. Bang, D.J. Dixon, and R.K. Sani. 2009. Isolation and characterization of cellulose-degrading bacteria from the deep subsurface of the Homestake gold mine, Lead, South Dakota, USA. *J. Ind. Microbiol. Biotechnol.* 36, 585-598.

Rastogi, G., R.K. Sani, B.M. Peyton, J.G. Moberly, and T.R. Ginn. 2008. Molecular studies on the microbial diversity associated with mining-impacted Coeur d' Alene river sediments. *Microb. Ecol.* DOI: 10.1007/s00248-008-9445-0.

Rosado, A.S., G.F. Duarte, L. Seldin, and J.D. Van Elsas. 1998. Genetic diversity of *nifH* gene sequences in *Paenibacillus azotofixans* strains and soil samples analyzed by denaturing gradient gel electrophoresis of PCR-amplified gene fragments. *Appl. Environ. Microbiol.* 64, 2770-2779.

Schloss, P.D. and J. Handelsman. 2005. Introducing DOTUR, a computer grogram for defining operational taxonomic units and estimating species richness. *Appl. Environ. Microbiol.* 71, 1501-1506.

Schloss, P.D. and J. Handelsman. 2006a. Introducing SONS, a tool for operational taxonomic unit-based comparisons of microbial community memberships and structures. *Appl. Environ. Microbiol.* 72, 6773-6779.

Schloss, P.D. and J. Handelsman. 2006b. Toward a census of bacteria in soil. *PLoS Comput. Biol.* e92.

Southham, G. and T.J. Beveridge. 1992. Enumeration of thiobacilli within pH-neutral and acidic mine tailings and their role in the development of secondary mineral soil. *Appl. Environ. Microbiol.* 58, 1904-1912.

Takai, K., D.P. Moser, M. DeFlaun, T.C. Onstott, and J.K. Fredrickson. 2001. Archaeal diversity in waters from deep South African gold mines. *Appl. Environ. Microbiol*. 67, 5750-5760.

Vanparys, B., K. Heylen, L. Lebbe, and P. De Vos. 2005. *Devosia limi* sp. nov., isolated from a nitrifying inoculum. *Int. J. Syst. Evol. Microbiol.* 55, 1997-2000.

Wang, Q., G.M. Garrity, J.M. Tiedje, and J.R. Cole. 2007. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* 73, 5261-5271.

Wang, C.M., C.L. Shyu, S.P. Ho, and S.H. Chiou. 2008. Characterization of a novel thermophilic, cellulose-degrading bacterium *Paenibacillus* sp. strain B39. *Lett. Appl. Microbiol.* 47, 46-53.