The Journal of Microbiology, February 2008, p. 34-39 Vol. 46, No. 1
DOI 10.1007/s12275-007-0247-8

A Novel Archaeal Group in the Phylum Crenarchaeota Found Unexpectedly in an Eukaryotic Survey in the Cariaco Basin

Sun-Ok Jeon[1], Tae-Seok Ahn[1], and Sun-Hee Hong[2]*

[1]*Department of Environmental Science, Kangwon National University, Chuncheon 200-701, Republic of Korea*
[2]*Department of Biology, Northeastern University, Boston, Massachusetts, 02115, USA*

(Received December 17, 2007 / Accepted January 25, 2008)

Archaea **have been found in many more diverse habitats than previously believed due in part to modern molecular approaches to discovering microbial diversity. We report here an unexpected expansion of the habitat diversity of the Archaea in the Cariaco Basin we found using a primer set designed for 18S eukaryotic rDNA sequence analysis. The results presented here expand the originally identified 9 archaeal clones reported in this environment using bacterial/archaeal primers to 152 archaeal clones: 67 (18 OTU) of these clones were found at a depth of 900 m of station A while 71 (9 OTU) of them were at a depth of between 300~335 m of station B&C depending upon which location the samples were taken. We used three phylogenetic analysis methods and detected 20 phylotypes belonging to a single previously unreported group distantly related to the Crenarchaeota. Also, we determined that the original nine sequences did not fall into any of the known phyla of the Archaea suggesting that they may represent a novel group within the Kingdom Archaea. Thus, from these two studies, we suggest that Archaea in the Cariaco Basin could be unique; however, further studies using archaeal-specific primers and the design of new primers as well as the systematic use of several different primer combinations may improve the chances of understanding the archeal diversity in the Cariaco Basin.**

Keywords**:** Archaeal diversity, Cariaco Basin, novel archaeal phylotypes

The Cariaco Basin, located in the north central coast of Venezuela (Fig. 1), and it is one of the world's largest anoxic marine water column system. There is no oxygen in this basin below depths of about 250 to 350 m (Madrid *et al*., 2001; Stoeck *et al*., 2003; Tedesco *et al*., 2007) because it is nearly 1,400 m deep with restricted circulation caused by a sill at 90 to 140 m. The basin's water column is characterized by a pronounced and predictable vertical layering of microbial communities and productivity. Significant increases in the abundance of prokaryotes, protozoa, viruses and prokaryotic productivity are observed routinely in the sulfidic waters immediately underlying the oxic/anoxic interface (Taylor *et al*., 2001, 2003).

Our ongoing study of the Cariaco Basin is aimed at investigating its eukaryotic diversity because we believed that the number of species in this environment could be more diverse and complex. Interestingly, we found many archaeal clones using a primer set typically designed for 18S eukaryotic rDNA sequence analysis during our study and the results presented here show significant archaeal diversity among the very poorly understood Archaea in this environment. To our knowledge, there have been two studies published (Madrid *et al*., 2001; Lin *et al*., 2006) on prokaryotic diversity focusing on bacterial rather than archaeal diversity which can provide the important ecological functions, such as sulfide oxidation, sulfate reduction, methanogenesis, and anaerobic oxidation of methane in the Cariaco Basin.

Archaea were originally divided into two phyla and described from extreme environments because the cultivation-based analysis led to the belief that they were restricted to extreme environments (Bano *et al*., 2004). Today, the Archaea domain can be divided into four phyla, Euryarchaeota, Crenarchaeota, Korarchaeota, the presence of which has been determined only by environmental DNA sequences (Barns *et al*., 1996; Bano *et al*., 2004), and reported Nanoarchaeota (Huber *et al*., 2002).

Furthermore, many studies based on 16S rRNA gene se-

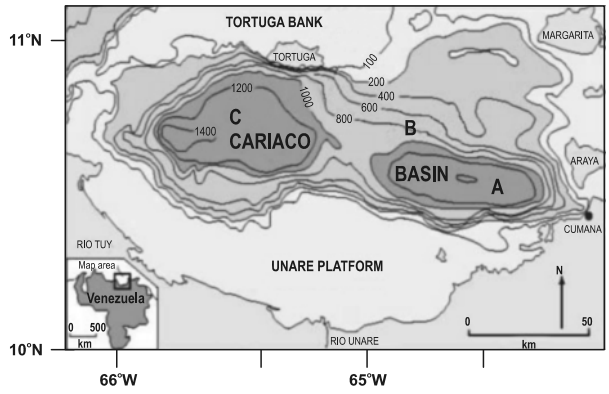

Fig. 1. Map of the Cariaco Basin sampling sites.

* To whom correspondence should be addressed.
(Tel) 1-617-373-3229; (Fax) 1-617-373-3724
(E-mail) s.hong@neu.edu

quence analyses, have demonstrated that Archaea are much more widespread than previously suspected since they have been found in a much wider variety of habitats (Takai and Horikoshi, 1999; Kim *et al*., 2000; Massana *et al*., 2000; Karner *et al*., 2001; Takai *et al*., 2001; Huang *et al*., 2002; Bano *et al*., 2004; Kim *et al*., 2005). Thus, molecular biological approaches for studying archaeal diversity have brought new perspectives to the study of microbial ecology and diversity, especially, during the past few years. Using these molecular techniques, we report here the results of a phylogenetic analysis of a collection of previously undescribed 16S rRNA gene sequences belonging to the domain Archaea.

Materials and Methods

Study area and sampling

Sampling was conducted in May 2004 aboard the B/O *Hermano Gines*, operated by Estacio'n de Investigaciones Marinas (EDIMAR) of the Fundacio'n La Salle de Ciencias Naturales, located on Margarita Island, Venezuela. Two or four liters of seawater were collected using a SeaBird rosette equipped with 12 TFE-lined, 8-L Niskin bottles under N_2 atmosphere to prevent oxygenation of samples and then evacuated 2-L hospital intravenous bags (non-DEHP vinyl; Secure Medical Inc., USA) immediately after retrieval. Samples were taken above the interface, at the interface, below the oxycline, and at 900 m depth from three different stations (Fig. 1) varying in their chemical and community gradients intensities and microbiological production. Duplicate bags from each depth were filtrated through 47 mm Durapore membranes (0.65-μm pore-size) (Millipore, USA) under gentle vacuum (<40 ml/min). Immediately after filtration, membranes were placed in individual 2 ml cryovials containing 1 ml of DNA extraction buffer (Zhou *et al*., 1996) with 5 μl of proteinase K and shipped frozen to the laboratory.

DNA isolation

Nucleic acids were extracted as described in Zhou *et al*. (1996) with slight modifications. Briefly, the samples were incubated for 30 min at 37°C in falcon tubes with 2 ml of a high-salt extraction buffer [100 mM Tris-HCl; pH 8.0, 100 mM Na_2EDTA; pH 8.0, 100 mM $NaPO_4$; pH 8.0, 1.5 M NaCl, and 1% cetyltrimethylammonium bromide (CTAB)] and 20 μl of proteinase K (final concentration, 100 μg/ml). After incubation, 200 μl of 20% (w/v) SDS was added, and heated them in a 65°C water bath for 2 h. These lysates were purified twice by extraction with an equal volume of chloroform-isoamyl alcohol (24:1, v/v) and were precipitated with 0.6 ml of isopropanol. The integrity of the total DNA was assessed on 0.8% agarose gels.

PCR amplification, cloning, and sequencing

Approximately 1,000 bp fragments were amplified using primer set, E528F-Univ1492RE (Edgcomb *et al*., 2002). The PCR protocol used HotStart *Taq* DNA polymerase (QIAGEN, USA), an initial hot-stat incubation for 15 min at 95°C, followed by 30 identical amplification cycles (denaturing at 95°C for 45 sec, annealing at 55°C for 1 min, and extension at 72°C for 2.5 min), and final extension at 72°C for 7 min. Each PCR product was cloned using TOPO TA cloning kit (Invitrogen, USA) as recommended by the manufacturer. To check the insert was of the correct length, colonies were screened by PCRs with primer set M13F; 5'-GTA AAA CGA GGC CAG-3' and M13R; 5'-CAG GAA ACA GCT ATG AC-3'. Clones were sequenced in 384-well plates using an ABI 3730XL capillary sequencer. Sequences editing, assembly of contigs, and BLAST analyses were performed using a series of UNIX shell scripts that incorporated PHRED and PHRAP (assembled by Hilary Morrison and Rich Fox at the Marine Biological Laboratory).

Phylogenetic analyses

The rRNA sequences were aligned by CLUSTAL X (Thompson *et al*., 1997). Potential chimeric sequences were screened by neighbor-joining trees constructed using PAUP (Swofford, 2001) and Chimera_Check available at the Ribosomal Database Project (RDP) II website (Maidak *et al*., 2001). The sequences were grouped into Operation Taxonomic Units (OTUs) based on a 97% rRNA gene sequence similarity level. This grouping was achieved by first making all possible pair-wise sequence alignments by using CLUSTAL W at default settings and calculating % sequence identities, followed by clustering the sequences into OTUs using PAUP (Swofford, 2001). Phylogenetic trees were constructed for the 152 archaeal clones using maximum parsimony, neighbor-joining and minimum evolution methods provided in Molecular Evolutionary Genetics Analysis 2.1 software (MEGA, version 4) (Tamura *et al*., 2007). Bootstrapping was used to estimate the reliability of phylogenetic reconstructions (1,000 replicates). The sequences determined in this study were submitted to the GenBank database, and are designated by the accession numbers EU369766-EU369917.

Results and Discussions

We originally carried out a molecular survey of eukaryotic diversity in the Cariaco Basin with four different eukaryotic primer sets: (i) E528F-Univ1391RE, (ii) E528F-Univ1492RE, (iii) E528F-Univ1517, and (iv) 360FE-Univ1492RE (Medlin *et al*., 1988; Lane, 1991; Shopsin *et al*., 1999; Edgecombe *et al*., 2002). Unexpectedly, we obtained 152 archaeal clones of the 1712 clones found at three different stations (A, B, and C) with primer set E528F and Univ1492RE; whereas, there were no significant archaeal clones with the other primer sets (less than 0.2%).

This is not surprising since even primers that bind to evolutionarily conserved regions of the SSU rRNA gene are never 100% universal at the domain level (Schmalenberger *et al*., 2001) and Wilms *et al*. (2006) found a novel archaeal cluster in tidal flats with primers that should amplify the first 500 bp for eukaryotic 18S rRNA genes. Also, Lovejoy *et al*. (2006) reported one or two archaeal sequences using the same 528F primer as ours using the RFLP method in sampling sites in the Arctic Ocean.

However, what is surprising is the high proportion (60%) of archaeal clones to total clones at Station A (900 m depth) (Table 1) given that there have been only 9 other archaeal clones found from the Cariaco Basin up to now. Despite the fact that Madrid *et al*. (2001) used universal

bacterial and archaeal primer sets 530F and 1494R to amplify either bacterial and archaeal 16S rRNA genes, they found one archaeal clone at Station A (500 m depth) and eight clones at Station A (1,310 m depth).

Thus, we believe that the 152 archaeal clones we found can be used as a starting point to investigate further the archaeal community structure in the Cariaco basin as well as in other variable environments because none of these clones have been detected previously in any natural environment.

In this study, 67 of the total archaeal clones (152) were found at 900 m at Station A while only 1 clone of the 152 clones was found at the same depth at Stations B&C. On

Table 1. Clones detected from the Cariaco Basin using primer set E528F, Univ1492RE

Sampling station	Sampling depths	The number of total clones	The number of archaeal clones	The ratio of archaeal clones to total clones (%)
Station A	Above[a]	339	0	0
	Interface[b]	67	1	1
	Below[c]	399	5	1
	900 m	111	67	60
Station B&C	Above	27	1	4
	Interface	68	6	9
	Below	274	71	26
	900 m	427	1	0
	Total	1,712	152	9

[a] Above: above of the oxic/anoxic interface (Station A=250 m; Station B=220 m; Station C=225 m)
[b] Interface: interface of the oxic/anoxic interface (Station A=280 m; Station B=245 m; Station C=280 m)
[c] Below: below of the oxic/anoxic (Station A=340 m; Station B=300 m; Station C=335 m)

Table 2. Frequency distribution of archaeal OTUs obtained in the Cariaco Basin

OTUs	The number of clones in library								Total no. of clones
	Station A				Station B&C				
	Above	Interface	Below	900 m	Above	Interface	Below	900 m	
1			3			1			4
2		1	3			3		1	8
3				11			7		18
4				1		1	11		13
5		2	1	28	1	4	36		72
6				2					2
7				2		5	10		17
8				3					3
9		2	1	3			4		10
10				3					3
11				3			3		6
12				1			1		2
13				1			1		2
14				2					2
15				1					1
16				1			1		2
17				1					1
18		1		1					2
19				1					1
20				1					1

[a] Above: above of the oxic/anoxic interface (Station A=250 m; Station B=220 m; Station C=225 m)
[b] Interface: interface of the oxic/anoxic interface (Station A=280 m; Station B=245 m; Station C=280 m)
[c] Below: below of the oxic/anoxic (Station A=340 m; Station B=300 m; Station C=335 m)

Fig. 2. Phylogenetic tree of representative novel archaeal phylotypes obtained from the Cariaco Basin. Trees were inferred by Maximum Parsimony (A), Neighbor-Joining (B), and Minimum Evolution (C). Bootstrap values are based on 1,000 replicates each and no values are given for groups with Bootstrap value less than 50%. The scale bar represents 0.05 (5%) nucleotide sequence difference. The GenBank accession numbers of compared sequences are as follows: X81886; *Pyrobaculum neutrophilum*, CP000505; *Thermofilum pendens*, CP000561; *Pyrobaculum calidifontis*, AB063645; *Vulcanisaeta souniana*, BA000002; *Aeropyrum pernix*, X99554; *Thermodiscus maritimus*, CP000493; *Hyperthermus butylicus*, X99560; *Staphylothermus marinus*, X99556; *Thermosphaera aggregans*, X98064; *Sulfophobococcus zilligii*, AY862079; Uncultured korarchaeote, AF255604; *Korarchaeota SRI-306*, AJ318041; *Nanoarchaeum equitans*, AB274308; *Archaeoglobus* sp., CP000300; *Methanococcoides burtonii*, AB220647; *Halococcus thailandensis*, AY350742; *Methanobacterium beijingense*, EF112196; *Methanobacterium* sp., CP000609; *Methanococcus maripaludis*. MY1106; *Pseudomonas* sp., AF052937; *Juncella racemosa*, NCMS601; *Ciliate* sp. The accession no. of AF224831, AF224869, AF224855, and AF224779 were from the original study in the Cariaco Basin (Madrid *et al.*, 2001). The accession no. from this study are as follows; OTU1:EU369884, OTU2:EU369854, OTU3:EU369865, OTU4:EU369839, OTU5:EU369769, OTU6:EU369852, OTU7:EU369902, OTU8:EU369907, OTU9:EU369914, OTU10:EU369916, OTU11:EU369910, OTU12:EU369897, OTU13:EU369899, OTU14:EU369886, OTU15:EU369870, OTU16:EU369873, OTU17:EU369875, OTU18:EU369909, OTU19:EU369869, OTU20:EU369867 (www.marscigrp.org/gsa05whl.ppt).

the other hand, 71 clones of the 152 total clones at station B&C (Table 1) were found at the "below" depth.

The CARIACO time series Station A (Fig. 1) was studied most intensively because of the wealth of supporting data generated there by other investigators. In the original eukaryotic study, Station B was selected because advective intrusions of oxidant-rich waters are believed to enter the Basin through the Canal de al Tortuga and move toward Station A (Astor *et al*., 2003) and Station C was selected because it is furthest from source waters with the weakest gradients and slowest microbial reaction rates. However, because the samples from Stations B&C in the original study were combined, it was necessary for us to use the combined samples here. While we found significant differences in the proportion of archaeal clones to total clones between Stations A and B&C depending upon the depth at which we sampled, no unique (singleton) clones were detected at Station B&C. However, we found 9 OTUs at Station A that did not occur at Station B&C. This suggests that the Station A in the Cariaco Basin is more diverse than the other stations sampled (Table 2).

Our phylogenetic analysis of the 152 archaeal clones identified 20 different phylotypes at the 97% similarity cut off value. The three phylogenetic analysis methods used consistently placed all 20 phylotypes into one group, which was distantly related to all of the other known Crenarchaeotal sequences. Different outgroups (Bacteria and Eukarya) did not change the phylogenetic location of Cariaco Basin sequences in our tree. Bootstrap values supported the inferred topology (61% in MP, 87% in NJ, 83% in ME; Fig. 2) and confirmed that the evolutionary distances are long enough to consider these Cariaco Basin archaeal sequences as a separate group, distinct from other Crenarchaeotal groups (Fig. 2). We also have examined whether or not these sequences could be retrieved using universal archaeal PCR primers, 915R and 958R (Teske and Sorensen, 2008), but no sequence match were founded.

However, four of the sequences (with lengths sufficient to assess) reported by Madrid *et al*. (2001) as falling into the Crenarchaeotal group didn't fall into that group based on any of our three phylogenetic analyses. In fact, these four did not fall into any known archaeal groups in our analyses. We suggest that these sequences might represent a novel group within the kingdom Archaea.

In conclusion, we have identified by standard methods of phylogenetic analyses 152 new sequences (20 phylotypes) from the Cariaco Basin demonstrating that there is significantly more archaeal diversity in the Basin than previously reported. While it is clear that these 152 sequences are truly novel archaeal sequences, future studies will be needed to determine if these sequences were detected fortuitously because of our use of these eukaryotic primers, or if we would have detected them also if we had used standard archaeal-specific primers.

Acknowledgements

This research was funded in part by a US National Science Foundation Microbial Observatory grant MCB-0348341. Sun-Ok Jeon was supported by the 2nd phase Brain Korea 21 Project in 2006. We thank to Dr. Slava Epstein for his supporting and we also grateful to Professor Edward Jarroll for editing the English usage in the manuscript.

References

Astor, Y.M., F. Müller-Karger, and M.I. Scranton. 2003. Seasonal and interannual variation in the hydrography of the Cariaco Basin: implications for basin ventilation. *Cont. Shelf Res*. 23, 125-144.

Bano, N., S. Ruffin, B. Ransom, and J.T. Hollibaugh. 2004. Phylogenetic composition of Arctic Ocean archaeal assemblages and comparison with Antarctic assemblages. *Appl. Environ. Microbiol*. 70, 781-789.

Barns, S.M., C.F. Delwiche, J.D. Palmer, and N.R. Pace. 1996. Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proc. Natl. Acad. Sci. USA* 93, 9188-9193.

Edgcomb, V.P., D.T. Kysela, A. Teske, A. De Vera Gomez, and M.L. Sogin. 2002. Benthic eukaryotic diversity in the Guaymas Basin hydrothermal vent environment. *Proc. Natl. Acad. Sci. USA* 99, 7658-7662.

Huang, L.N., H. Zhou, Y.Q. Chen, S. Luo, C.Y. Lan, and L.H. Qu. 2002. Diversity and structure of the archaeal community in the leachate of a full-scale recirculating landfill as examined by direct 16S rRNA gene sequence retrieval. *FEMS Microbiol. Lett.* 214, 235-240.

Huber, H., M.J. Hohn, R. Rachel, T. Fuchs, V.C. Wimmer, and K.O. Stetter. 2002. A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* 417, 63-67.

Karner, M.B., E.F. DeLong, and D.M. Karl. 2001. Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* 409, 507-510.

Kim, B.S., H.M. Oh, H.J. Kang, and J.S. Chun. 2005. Archaeal diversity in tidal flat sediment as revealed by 16S rDNA analysis. *J. Microbiol*. 43, 144-151.

Kim, H., D. Honda, S. Hanada, and N. Kanamori. 2000. A deeply branched novel phylotype found in Japanese paddy soils. *Microbiology* 146, 2309-2315.

Lane, D.J. 1991. 16S/23S sequencing, p. 115-175. *In* E. Stackebrandt and M. Goodfellow (eds.), Nucleic acid technologies in bacterial systematic. John Wiley and Sons, Chichester, UK.

Lin, X., S.G. Wakeham, I.F. Putnam, Y.M. Astor, M.I. Scranton, A.Y. Chistoserdov, and G.T. Taylor. 2006. Comparison of Vertical distributions of prokaryotic assemblages in the anoxic Cariaco Basin and Black Sea by use of fluorescence in situ hybridization. *Appl. Environ. Microbiol*. 72, 2679-2690.

Lovejoy, C., R. Massana, and C. Pedrós-Alió. 2006. Diversity and distribution of marine microbial eukaryotes in the arctic ocean and adjacent seas. *Appl. Environ. Microbiol.* 72, 3085-3095.

Madrid, V.M., G.T. Taylor, M.I. Scranton, and A.Y. Chistoserdov. 2001. Phylogenetic diversity of bacterial and archaeal communities in the anoxic zone of the Cariaco Basin. *Appl. Environ. Microbiol*. 67, 1663-1674.

Maidak, B.L., J.R. Cole, T.G. Lilburn, C.T. Parker, Jr., P.R. Saxman, R.J. Farris, G.M. Garrity, G.J. Olsen, T.M. Schmidt, and J.M. Tiedje. 2001. The RDP-II (Ribosomal Database Project). *Nucleic Acids Res*. 29, 173-174.

Massana, R., E.F. DeLong, and C. Pedros-Alio. 2000. A few cosmopolitan phylotypes dominate planktonic archaeal assemblages in widely different oceanic provinces. *Appl. Environ. Microbiol*. 66, 1777-1787.

Medlin, L., H.J. Elwood, S. Stickel, and M.L. Sogin. 1988. The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. *Gene* 71, 491-499.

Schmalenber, A., F. Schwieger, and C.C. Tebbe. 2001. Effect of pri-

mers hybridizing to different evolutionarily conserved regions of the small-subunit rRNA gene in PCR-based microbial community analyses and genetic profiling. *Appl. Environ. Microbiol*. 67, 3557-3563.

Shopsin, B., M. Gomez, S.O. Montgomery, D.H. Smith, M. Waddington, D.E. Dodge, D.A. Bost, M. Riehman, S. Naidich, and B.N. Kreiswirth. 1999. Evaluation of protein A gene polymorphic region DNA sequencing for typing of *Staphylococcus aureus* strains. *J. Clin. Microbiol*. 37, 3556-3563.

Stoeck, T., G.T. Taylor, Y. Astor, and S. Epstein. 2003. Novel eukaryotes from a permanently anoxic Cariaco Basin (Caribbean Sea). *Appl. Environ. Microbiol*. 69, 5656-5663.

Swofford, D.L. 2001. PAUP*. Phylogenetic analysis using parsimony (and other methods), 4.0 beta 10 ed. Sinauer Associates, Sunderland, MA, USA.

Takai, K. and K. Horikoshi. 1999. Genetic diversity of Archaea in deep-sea hydrothermal vent environments. *Genetics* 152, 1285-1297.

Takai, K., D.P. Mpser, 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.

Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007. MEGA 4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24, 1596-1599.

Taylor, G.T., M. Scranton, M. Iabichella, T.Y. Ho, and R. Varela. 2001. Chemoautotrophy in the redox transition zone of the Cariaco Basin, a significant source of midwater organic carbon production. *Limnol. Oceanogr*. 46, 148-163.

Taylor, G.T., C. Hein, and M. Iabichella. 2003. Temporal variations in viral distributions in the anoxic Cariaco Basin. *Aquatic Microbiol. Ecol*. 30, 103-116.

Tedesco, T., R. Thunell, Y. Astor, and F.M. Karger. 2007. The oxygen isotope composition of planktonic foraminifera from the Cariaco Basin, Venezuela: Seasonal and interannual variations. *Marine Micropaleontology* 62, 180-193.

Teske, A. and K.B. Sorensen. 2007. Uncultured archaea in deep marine subsurface sediments: have we caught them all? *ISME J.* 2, 3-18.

Thompson, J.D., T.J. Gibson, F. Plewniak, F. Jeanmougin, and D.G. Higgins. 1997. CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*. 24, 4876-4882.

Wilms, R., H. Sass, B. Köpke, J. Köster, H. Cypionka, and B. Engelen. 2006. Specific bacterial, archaeal, and eukaryotic communities in tidal-flat sediments along a vertical profile of several meters. *Appl. Environ. Microbiol*. 72, 2756-2764.

Zhou, J., M.A. Bruns, and J.M. Tiedje. 1996. DNA recovery from soils of diverse composition. *Appl. Environ. Microbiol*. 62, 316-322.