Phylogenetic analysis of Group I marine archaeal rRNA sequences emphasizes the hidden diversity within the primary group Archaea

JAMES O. MCINERNEY[1], MARICE MULLARKEY[2], MARTINA E. WERNECKE[2] AND RICHARD POWELL[2*]

[1]*Department of Zoology, The Natural History Museum, London SW7 5BD, UK*
[2]*Department of Microbiology, University College Galway, Ireland* (richard.powell@ucg.ie)

SUMMARY

Archaea form one of the three primary groups of extant life and are commonly associated with the extreme environments which many of their members inhabit. Currently, the Archaea are classified into two kingdoms, Crenarchaeota and Euryarchaeota, based on phylogenetic analysis of ribosomal RNA (rRNA) sequences. Molecular techniques allowing the retrieval and analysis of rRNA sequences from diverse environments are increasing our knowledge of archaeal diversity. This report describes the presence of marine Archaea in north-east Atlantic waters. Quantitative estimates indicated that the marine Archaea constitute 8% of the total prokaryotic rRNA in Irish coastal waters. Phylogenetic analysis of the archaeal rRNA gene sequences revealed sufficient genetic diversity within Archaea to indicate that the current two-kingdom classification of Crenarchaeota and Euryarchaeota is restrictive.

1. INTRODUCTION

Molecular methods have added great impetus to microbial systematics and ecology, and the need for this impetus is clear as estimates that less than 5% of micro-organisms are known (Atlas *et al*. 1992) mean that vast microbial resources remain concealed. With current estimates for the number of extant bacterial species approaching three million (Watson *et al*. 1995), one can only speculate that many major groups of bacteria are yet undiscovered. One group of organisms recognized through molecular studies, the Archaea (Archaebacteria) (Balch *et al*. 1977; Woese & Fox 1977), comprise an evolutionary related but ecologically diverse group of extreme thermophiles (kingdom Crenarchaeota), and thermophiles, sulphur metabolizers, halophiles and methanogens (kingdom Euryarchaeota) (Woese *et al*. 1990). Commonly termed extremophiles, the Archaea are known for the inhospitable environments they inhabit, including hot springs, deep-sea hydrothermal vents and high-saline waters. These unusual habitats have contributed to the notion that Archaea are non-competitive relics with little importance in global ecology or biodiversity. This picture began to change in 1992, when two concurrent reports revealed the unexpected presence of Archaea among marine bacterioplankton in north-east Pacific (DeLong 1992; Fuhrman *et al*. 1992) and north-west Atlantic (DeLong 1992) ocean waters. In both cases, the presence of Archaea was indicated by molecular techniques retrieving small subunit rRNA gene sequences (SSU rDNA) using the polymerase chain reaction (PCR). Phylogenetic analysis of the recovered archaeal SSU rDNA sequences divided the marine Archaea into two groups and indicated that Group I constituted a new deep-branching lineage within the Crenarchaeota, and Group II represented a novel lineage within the Euryarchaeota (DeLong 1992). Subsequently, related archaeal Group I SSU rDNA sequences have been recovered from Antarctic coastal waters (DeLong *et al*. 1994), the gut contents of abyssal holothurians (McInerney *et al*. 1995), and in symbiosis with a marine sponge (Preston *et al*. 1996). To date, no members of either group of the marine Archaea have been cultivated and we have little idea of their physiology or ecological role. However, their potential importance within marine ecosystems is reflected by Northern blot analysis which revealed these micro-organisms to represent approximately 2% and 30% of prokaryotic SSU rRNA in Pacific coastal (DeLong 1992) and Antarctic (DeLong *et al*. 1994) waters, respectively.

Here, we report the recovery and phylogenetic analysis of further marine archaeal Group I SSU rDNA sequences from a series of sites and depths in the north-east Atlantic ocean. Near full-length archaeal SSU rDNA sequences were amplified by PCR from DNA extracted from ocean water samples using

*Author for correspondence.

primers designed to selectively recognize Archaea. The SSU rDNA amplification products were subcloned into *E. coli* and the nucleotide sequences were determined. The archaeal SSU rDNA sequences were aligned to and compared with a broad database of rRNA sequences including cultivated and non-cultivated members of the Archaea, Bacteria, and Eucarya to determine their higher-order phylogenetic relationships. Their relationship with other reported marine archaeal SSU rDNA sequences was also examined. Finally, marine Archaea activity was estimated at one site by Northern blot analysis of total prokaryotic SSU rRNA. This report describes an increasing ecological range for the marine Archaea in ocean waters with a correlation between phylotype and geographical location for many of the retrieved SSU rDNA sequences. Phylogenetic analysis of the marine archaeal sequences and SSU rDNA sequences from cultivated and uncultivated Archaea revealed sufficient genetic diversity to suggest that the current two-kingdom classification of Crenarchaeota and Euryarchaeota understates the true diversity of extant Archaea.

2. MATERIALS AND METHODS

Ocean water samples were collected from two different locations; 20 l samples were collected on board the Russian research vessel R. V. Prof. Marti at positions 53° N 14°58′ W (500 m), 55° N 11° W (500 m), 56° N 11°30′ W (500 m), 59° N 9° W (0 m), 60° N 5°30′ W (100 m); and 80 l samples were collected 200 m offshore at the Carna research station, at position 53°19′ N 9°50′ W (10 m). The samples were prefiltered through a 10-μm mesh filter at the Carna site, and through a 2-μm GF/C filter (Whatman Int. Ltd, UK) on the R. V. Prof. Marti. Samples were concentrated to approximately 200 ml using a tangential flow filtration device fitted with a filter of nominal pore size 0.2 μm (Sartorius AG, Germany). Concentrated water samples collected on the R. V. Prof. Marti were filtered onto 47 mm diameter 0.2 μm pore size cellulose acetate filters (Sartorius AG). DNA was extracted by a previously described method (Fuhrman *et al*. 1988). Concentrated seawater samples obtained from the Carna coastal site were treated using a direct lysis procedure. 100 ml of the sample were centrifuged at 8000 × *g* for 5 min, the supernatant was discarded and the pellet was resuspended in 2 ml of 0.1 M EDTA (pH 8.0). 1.0 ml of 10 mg ml^{-1} lysozyme (Boehringer Mannheim, Germany) was added, and the sample was incubated at 37 °C for 1 h, followed by the addition of 200 μl of 10 mg ml^{-1} proteinase K (Boehringer Mannheim) after which the samples were held at room temperature for 10 min. 500 μl of 20% SDS were then added and the samples were left at room temperature for 15 min with gentle agitation. 2 ml of 0.1 M EDTA (pH 8.0) were added, and the solution was gently mixed. An equal volume of phenol was added, followed by a gentle swirling of the samples for 1 min and centrifugation at 10 000 × *g* for 10 min. Two successive extractions using phenol : chloroform : isoamyl alcohol (24 : 24 : 1) were performed. The same volume was used in each. The recovered aqueous phase was twice extracted using the same volume of chloroform : isoamyl alcohol (24 : 1) each time, and DNA was precipitated in 4 volumes of ethanol. The amplification of archaeal SSU rDNA sequences was perfomed in a similar manner to that which we have previously described (McInerney *et al*. 1995), without the second round of nested amplifications. The designations and sequences of the PCR primers and their reference positions on alignments of SSU rRNAs are AB (Archaea specific), 5′-TCCGGTTGATCCTGCCGG-3′ (bases 3 to 21); and UN (universal), 5′-ACGGNWACCTTGTTACGAGTT-3′ (bases 1423 to 1402) (standard International Union of Pure and Applied Chemistry nomenclature: N is G, A, T, or C; W is A or T). The 100 μl amplification reactions contained one-tenth volume 10 × PCR buffer (Promega, WI; 200 mM TrisHCl pH 8.5, 15 mM $MgCl_2$, 500 mM KCl, 1% tween), 120 ng of each primer, 20 ng of target DNA, and 1.0 U of Taq polymerase (Promega) overlayed with 30 μl of mineral oil. The PCR cycling conditions were: 30 cycles of 95 °C for 1 min, 55 °C for 1 min, and 72 °C for 2 min, and were performed using the TRIO-thermoblock (Biometra GmbH, Germany).

Five amplification products from a single Prof. Marti site were pooled and, along with the Carna site products, were purified using the Magic PCR Preps DNA Purification system (Promega) and cloned into *E. coli* cells using the PCR-Script SK(+) (Stratagene, CA) according to the manufacturers instructions. White colonies were screened for the presence of PCR inserts by hybridization with 100 ng of PCR products radiolabelled with 50 μCi (α-^{32}P) dCTP (DuPont NEN, UK) using the rediprime system (Amersham, UK). DNA sequencing was carried out in a stepwise manner using plasmid forward and reverse primers and a number of primers based on conserved regions of Archaea and Bacteria SSU rRNA gene sequences (Lane 1991). The sequences reported in this paper have been deposited in the GenBank database under the accession numbers U71109–U71118.

A representative alignment of prokaryotic and eukaryotic taxa was retrieved from the Ribosomal Database Project (RDP) (Olsen *et al*. 1991) and the sequences from this study were added. All alignments were carried out by eye using the Genetic Data Environment (GDE 2.2) program (Smith 1994). A total of ten regions of ambiguous alignment were removed and the remaining regions were concatenated together. The initial alignment included 134 taxa and was 3229 positions in length. Preliminary phylogenetic analyses were carried out on this alignment to identify sequences from the major groupings. When multiple sequences were available from a particular phylogenetic grouping, the sequence with the most balanced base composition was chosen. Maximum parsimony searches were carried out heuristically. Random addition of sequences was carried out ten times. Distance matrices were calculated by the LogDet (Lockhart *et al*. 1994) and gg95 (Galtier & Gouy 1995) methods and the results of these calculations were used to construct phylogenetic hypotheses using the neighbour-joining algorithm (Saitou & Nei 1987). The maximum likelihood model

that was used was that described by Felsenstein (Felsenstein 1981). The Kishino–Hasegawa (Kishino & Hasegawa 1989) test was carried out using constraint trees prepared by MacClade 3.01 (Maddison & Maddison 1992). Maximum likelihood analysis (including the Kishino–Hasegawa test) was carried out using fastDNAml 1.1.1 (Olsen *et al*. 1994). The phylo_win program of Nicolas Galtier (available from `ftp://biom3.univ-lyon1.fr/`) was used for calculating LogDet and gg95 distances and constructing neighbour-joining trees from these distances. AutoDecay 2.9.2 (Eriksson 1996) was used for creation of constraint files for use with PAUP 3.1.1 (Swofford 1993).

For RNA extraction, concentrated water samples were filtered onto 47 mm diameter, 0.2 μm pore size cellulose acetate filters (Sartorius AG). The filters were cut into strips (1 cm × 2 mm) and placed in 4 ml of STE buffer (100 mM NaCl, 10 mM TrisHCl, 1 mM EDTA, pH 8.0). 0.4 ml of 10% SDS were added and the tubes were placed in a boiling water bath for 2 min. The samples were centrifuged at 10 000 × *g* for 10 min at 15 °C and split into two 2.5 ml aliquots. 0.83 ml of 3 M sodium acetate (pH 5.2) were added, followed by 5 ml of ice-cold absolute ethanol. RNA was precipitated at −70 °C for 1 h. The sample was centrifuged at 10 000 × *g* for 30 min at 4 °C. The pellet was resuspended in 1 ml of lysis buffer (80 mM TrisHCl pH 7.5, 10 mM $MgCl_2$-$7H_2O$, 10 mM 2-mercaptoethanol). 100 μl of 2 mg ml^{-1} heparin (Boehringer Mannheim), and 100 μl of 2.3 mg ml^{-1} proteinase K (Boehringer Mannheim) and 100 μl of 0.1 M EDTA pH 8.0 were added followed by incubation at −70 °C for 1 h. The microfuge tube was then placed at 60 °C until the solution thawed, and was then transferred to a tube containing a heated mixture of 500 μl of phenol, 500 μl of chloroform : isoamyl alcohol (24 : 1), and 100 μl of 10% SDS, and incubated at 60 °C for 2 min with shaking. The tube was then kept on ice for 2 min before centrifugation at 7000 × *g* for 5 min at 4 °C. RNA from the aqueous phase was precipitated by the addition of one-tenth volume of 2.5 M sodium acetate and 2 volumes of absolute ethanol. The RNA was pelleted by centrifugation at 14 000 × *g* for 20 min and resuspended in 30 μl of sterile H_2O. Total RNA was fractionated by electrophoresis in denaturing 1% agarose (Bio/Gene, UK) gels (20 × 20 cm) prepared in 5% formaldehyde and 1 × MOPS buffer using a protocol similar to that previously described (Ausubel *et al*. 1992). 0.5 μg of total RNA prepared from *E. coli* and *Haloferax volcanii* and 0.3 μg of total RNA isolated from the Carna site were analysed. RNA was visualized by staining the gel with 1 μg ml^{-1} ethidium bromide, before viewing over UV. The gel was blotted onto a nytran membrane (Schleicher & Schuell, Germany) by capillary action. Following transfer the blot was air dried, and the RNA was fixed to the membrane by baking at 80 °C for 30 min. Hybridization was performed with an Archaea-specific SSU rRNA-targeted oligodeoxynucleotide DNA probe (Raskin *et al*. 1994), 5′-GTGCTCCCCCGCCAATTCCT-3′ (rRNA bases 934915) end-labelled with 50 μCi (γ-^{32}P) ATP (DuPont) and 1 U T4 polynucleotide kinase (Promega). Standard hybridization (6 × SSC, 5 × Denhardt's solution, 0.1% SDS) was performed for 8 h at 56 °C. The filter was then washed in 6 × SSC, 0.1% SDS at room temperature for 20 min, in 6 × SSC, 0.1% SDS at 40 °C for 20 min, in 1 × SSC, 0.1% SDS at 40 °C for 20 min, and in 0.1 × SSC, 0.1% SDS at 40 °C for 20 min, before autoradiography for 15 h. Densitometry analysis was performed using the Bio-Gene software package (Vilber Lourmat, France) on the Bio-Profil imaging system (Vilber Lourmat).

3. RESULTS

The north-east Atlantic waters analysed in this study included five samples taken from sites located between 200 and 400 km from the north-west Irish and Scottish coasts (Prof. Marti samples) spanning a latitudinal range of approximately 800 km, and one sample taken 200 m from the west Irish coast (Carna samples). The depths for the samplings ranged from surface waters to 500 m, and the sample volume ranged from 20–80 l. After concentration, total DNA was extracted from the samples and almost full-length archaeal SSU rRNA genes were amplified by PCR. The PCR products of the pooled Prof. Marti samples, and the Carna samples were subcloned into *E. coli* cells generating SSU rRNA gene libraries containing 250 clones and 50 clones, respectively. These SSU rDNA libraries were screened with a DNA probe representing the SSU rDNA sequences that had been amplified from each respective site to identify the more abundant archaeal SSU rDNA sequences contained within the libraries. 28% of the Prof. Marti clones (70 clones) and 30% of the Carna clones (15 clones) hybridized with these DNA probes. Three and seven clones were chosen at random for DNA sequencing from the Prof. Marti and Carna SSU rDNA libraries respectively. All ten SSU rDNA sequences were archaeal in origin and no evidence of chimeric PCR artefacts were noted.

Phylogenetic analysis was performed to determine the higher-order relationships of the retrieved marine archaeal sequences. Only positions that had homologues in all sequences were included and the alignment was 980 positions in length. A total of 52 equally parsimonious trees, each requiring 3291 steps, were found from a maximum parsimony search. In all ten replicates, minimal length trees were found. This is an indication that the search landscape is relatively smooth with no non-optimal search islands detected. A strict consensus tree could be produced that affected a total of ten taxa. These taxa were within the Carna clade (four taxa affected, yielding a polytomy) and within the cultured Crenarchaeota (six taxa affected, a trichotomy of *Desulfurococcus mobilis*, *Sulfolobus shibitae* and *Pyrodictium occultum*, and a polytomy at the base of this clade). All other partitions in the analysis were the same in all 52 trees. This analysis showed the marine Archaea Group I sequences occupying a position as a sister clade to the cultured Crenarchaeota. The LogDet (Lockhart *et al*. 1994) and gg95 (Galtier & Gouy 1995) distance matrix methods were also used, both of which are designed to reconstruct phylogenetic relationships from sequences that have evolved under a non-stationary evolutionary

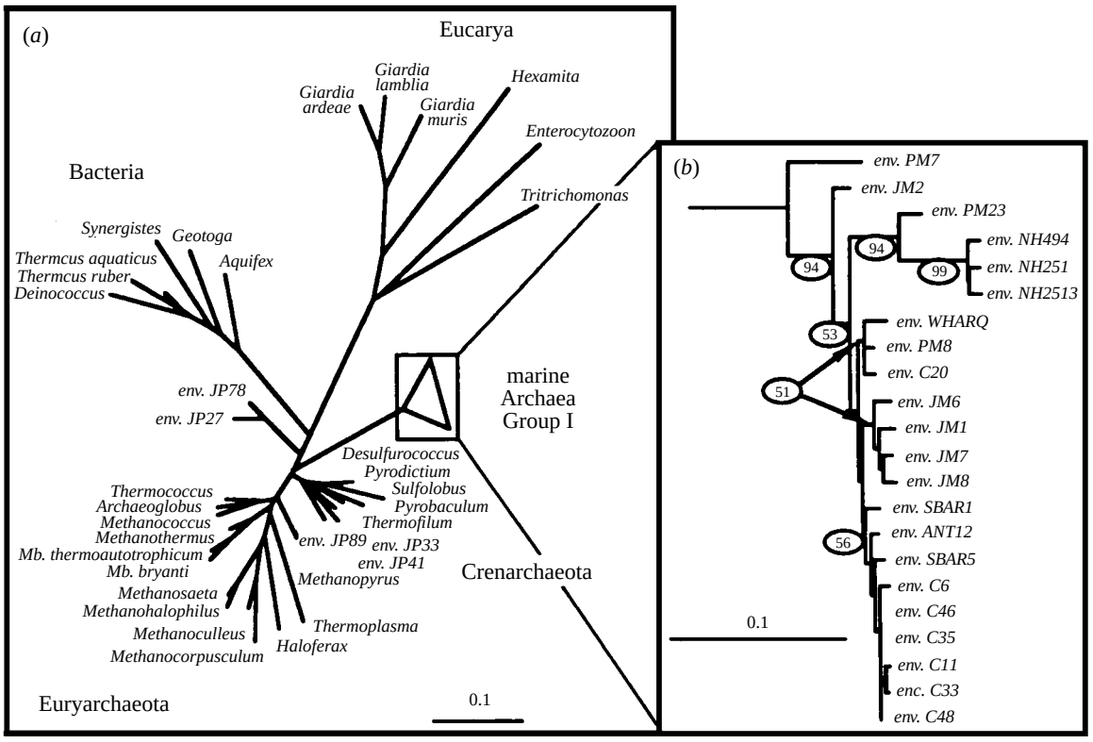

Figure 1. (*a*) Majority-rule consensus neighbour-joining tree constructed from LogDet distances of parsimony-informative sites. (*b*) Portion of the phylogenetic tree showing the internal relationships among the marine Archaea Group I SSU rDNA sequences. The bootstrap support by LogDet neighbour-joining analysis is shown by encircled numbers. The scale bar denotes an evolutionary distance of 10% sequence change.

process. The LogDet method was employed on parsimony-informative sites only, in accordance with the original description of the use of this method. In both cases, the phylogenetic trees placed the marine Archaea Group I sequences as a sister taxa to the Crenarchaeota (figure 1*a*). For maximum likelihood analyses the transition/transversion ratio for the data set was determined by an iterative process, with the maximum likelihood value being estimated at 1.0. The base frequency data for the analysis was determined empirically from the data set. The maximum likelihood tree also placed the marine Archaea Group I as a sister taxa to the cultured Crenarchaeota.

The robustness of the phylogenetic hypotheses were evaluated using bootstrap resampling (Felsenstein 1985) and Bremer support (decay) indices (Bremer 1988). The results for some of the more relevant clades are shown in figure 2. All of the sequences from this study formed a single group along with marine Archaea Group 1 SSU rDNA sequences previously described (figure 2A). No phylogenetic analysis method broke this grouping for any bootstrap replicates and under the maximum parsimony criteria, a total of 72 steps were required to break this clade (figure 2A). All of the methods place the marine archaeal sequences as a sister taxa to the Crenarchaeota. However, bootstrap analysis was not very strongly supportive of this placement, with parsimony and the gg95 method giving bootstrap proportions of 74 and 72 respectively, while the LogDet method only supported this placement 57% of the time (figure 2B). Also, a constrained analysis on the original data set required only two additional steps in order to break this clade. Two additional placements of the marine archaeal SSU rDNA Group I sequences were examined. The hypothesis that these sequences are sister taxa to the Euryarchaeota was not strongly favoured, appearing in only 8.58% of reconstructed trees from the bootstrap replicates (figure 2C). There was no data available for the frequency of this partition for either the gg95 or LogDet methods as this option is not available in the software. Also, there is no Bremer support value for this hypothesis as it did not appear on any of the original most parsimonious trees. Placement of the marine archaeal Group I sequences as an early separation from an ancestor of the cultured members of the Archaea also received little support (figure 2D). Using the maximum likelihood framework and the test devised by Kishino & Hasegawa (1989), the hypothesis placing the marine Archaea Group I as a deeper division than the Crenarchaeota–Euryarchaeota split was examined. The Kishino–Hasegawa significance test was performed at the 95% confidence level and the analysis shows this hypothesis is not significantly worse than the maximum likelihood hypothesis.

To determine the internal relationships within the uncultured marine Archaea Group I, an alignment of 407 positions was used. This included all the positions that were common to all the sequences of the marine Archaea Group I and allowed a significant expansion

Topology	Pars	gg95	LogDet	Bremer
marine Archaea Group I — A.	100	100	100	72
Crenarchaeota, marine Archaea Group I — B.	74	72	57	2
Euryarchaeota, marine Archaea Group I — C.	8.58	N/A	N/A	N/A
Euryarchaeota, Crenarchaeota, marine Archaea Group I — D.	5.05	N/A	21	N/A

Figure 2. Supports for the major partitions. Column 1 (topology) shows the relationships that were tested with the arrows pointing to the internal edge whose integrity is in question. Columns 2–4 show the values obtained by bootstrap resampling of the data set using maximum parsimony (Pars), gg95 and LogDet methods respectively. Column 5 shows the Bremer support value for the clades on either side of the internal branch. This value is found by the difference in tree length by parsimony analysis between the most parsimonious tree and the shortest tree that does not contain the clade of interest. In the case of the gg95 and LogDet distance methods, the entries marked n/a means that these data were unavailable from the phylo.win program. For the Bremer support analysis, n/a denotes values that were not available due to the clade of interest not being found on any of the 52 equally most parsimonious trees.

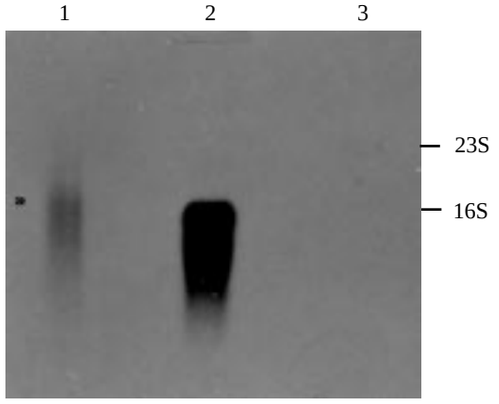

Figure 3. Northern analysis of prokaryotic SSU rRNA at the Carna site hybridized with an Archaea-specific DNA probe. Lane 1, total RNA from the Carna site; lane 2, total RNA from *H. volcanii* (Archaea); lane 3, total RNA from *E. coli* (bacteria). 16S and 23S rRNA size markers are indicated on the right. The arrow denotes hybridization with Archaeal SSU rRNA at the Carna site.

of the number of taxa of interest that could be used in the analyses. The data set was analysed in a number of ways, but the results of each method were similar. Figure 1*b* shows the internal relationships of the marine archaeal Group I clade. Three groupings were identified in the clade that were quite strongly supported. Three sequences (env.NH494, env.NH251, env.NH2513) obtained from Pacific Ocean bacterioplankton (Fuhrman *et al*. 1992) were seen to cluster together 99% of the time, and the env.PM23 sequence was placed as a sister taxon in 94% of bootstrap replicates. The single sequence env.PM7 was separated from the rest of the sequences with a base pair value of 94. Among the rest of the taxa, the support for any particular partition is not very strong. An analysis of the bootstrap partition table reveals this to stem largely from the tendency of some taxa to change position on the tree, thereby having a destabilizing influence. However, there is an apparent correlation between the majority of the sequences and the studies from whence they came, i.e. the env.C6/C46/35/C11/C33/C48 sequences are from the Carna site of this study, the env.JM6/JM1/JM7/JM8 sequences were isolated from the gut contents of abyssal holothurians in the deep Atlantic (McInerney *et al*. 1995), and the env.NH494/251/2513 sequences were recovered from a Pacific ocean site. This evidence suggests that members of the marine Archaea Group I have evolved and adapted to particular ecological niches in ocean waters.

The relative activity of the marine Archaea at the Carna site was analysed by determining the relative proportion of total prokaryotic SSU rRNA that corresponded to members of this domain. Total SSU rRNA was prepared, and electrophoresed along with a selected candidate of the Bacteria, *E. coli*, and the Archaea, *Haloferax volcanii*, and analysed by Northern blot using an archaeal-specific SSU rRNA-directed oligodeoxynucleotide DNA probe. Hybridization could clearly be seen with the SSU rRNA of the archaeal candidate *H. volcanii* (figure 3, lane 2) while no hybridization was apparent with the bacterium, *E. coli* (figure 3, lane 3). Hybridization was also seen with the total SSU rRNA isolated from the Carna site (figure 3, lane 1, arrowed). Based on the known concentrations of total RNA examined along with densitometric comparison of the hybridization signals with the *H. volcanii* and Carna RNAs, we estimated that Archaea constitute approximately 8% of the prokaryotic SSU rRNA at the Carna site at the time of sampling.

4. DISCUSSION

The Archaea are currently segregated into two major lineages, kingdoms Crenarchaeota and Euryarchaeota, whose cultivated members are associated with unusual or extreme habitats. However, recent molecular studies have identified the presence of Archaea in general terrestrial and aquatic environments including soils from a soyabean field (Ueda *et al*. 1995) and boreal forest (Jurgens *et al*. 1997), and lake water sediments (MacGregor *et al*. 1997; Schleper *et al*. 1997). Currently, the marine Archaea have been identified in the coastal waters (0–500 m depths) of North America (DeLong 1992; Fuhrman *et al*. 1992) and Antarctica (DeLong *et al*. 1994) and associated with abyssal holothurian species in the deep Atlantic (McInerney *et al*. 1995). The origins of these marine Archaea are unknown. However, G : C ratios (which are less than those of

hyperthermophiles), distribution patterns, and rRNA hybridization analysis, suggest that some of these Archaea may represent uncultivated pscychrophilic members of bacterioplankton. This study supports this observation with the presence of archaeal SSU rDNA identified at six north-east Atlantic ocean sites ranging from 200 m to 400 km offshore over a latitude distance of approximately 800 km, and at depths from surface waters to 500 m. Also, Northern blot analysis at the Carna site revealed Archaea to constitute approximately 8% of the prokaryotic SSU rRNA which is similar to the previously reported 4% from the American Pacific coast (DeLong 1992), and less than the 21–34% reported for Antarctic waters (DeLong *et al*. 1994). The surface water temperature at the Carna site ranges annually from 7 °C to 16 °C and the Northern blot analysis was performed on a sample collected in January 1996 when the water temperature was approximately 7 °C. Similar levels of archaeal SSU rRNA were recorded at this site in March and May 1996. These marine Archaea contradict many of our previous conceptions of archaeal ecology, and they appear to represent a significant fraction of the bacterioplankton of cold ocean waters, one of the world's largest (by volume) habitats.

In terms of the higher-order relationships, the placement of the marine Archaea Group I as a sister kingdom to the Crenarchaeota was a common occurrence in all the phylogenetic analysis methods employed. However, the levels of support for this position do not warrant confidence. There are two explanations for these results. Either a rapid series of separation events occurred at the time when the common ancestors of the Crenarchaeota, Euryarchaeota and the marine Archaea diverged, or there is insufficient information within the SSU rRNA gene to resolve this issue. Interestingly, two environmental archaeal SSU rDNA sequences (env.pJP27 & env.pJP78) recovered from sediment of a hot spring (Barns *et al*. 1994) are placed in our analysis before the Crenarchaeota–Euryarchaeota separation (figure 1*a*). These, and other related, SSU rDNA sequences have recently been proposed as representatives of a new archaeal kingdom provisionally named Korarchaeota (Barns *et al*. 1996). Our independent analysis confirms the striking phylogeny of this group. The peripheral relationships of the marine Archaea Group I and env.pJP27/78 sequences to the kingdom Crenarchaeota suggest they may truely represent new archaeal kingdoms.

With respect to the internal relationships between the marine archaeal SSU rDNA sequences of this study and those reported previously, every method of phylogenetic analysis placed all the marine Archaea Group I sequences as a well-supported clade. It is not possible to directly ascribe species-level or genus-level differences between the organisms of this group on the basis of SSU-rRNA gene sequence alone. However, by analogy to SSU rRNA sequence differences among cultured Bacteria and Archaea, it is apparent that many of these taxa would be ascribed to different genera. The level of sequence dissimilarity among the marine Archaea Group I ranged from 10.5% (env.PM7 & env.JM1) to 1% (env.C35 & env.C48). The order Cenarchaeales and family name Cenarchaeceae have recently been proposed for a sponge-associated symbiont marine archaeal SSU rDNA sequence (Preston *et al*. 1996) and its planktonic (Delong 1992; DeLong *et al*. 1994; Fuhrman *et al*. 1992) and abyssal relatives (McInerney *et al*. 1995). All the sequences reported in this study would be members of this family. However, our phylogenetic analysis suggests that the assumption that the marine Archaea Group I are more recent adaptations of hyperthermophilic Crenarchaeota may be incorrect. Taking into account the placement of the env.pJP27/78 sequences, the marine Archaea Group I may well have adapted to cold ocean waters from thermophilic environments (i.e. their ancestor was thermophilic) but their membership of the kingdom Crenarchaeota is questionable. Indeed, our current knowledge of the kingdom Crenarchaeota based on cultivated representatives sharing the signature thermophilia, may be in need of revision in light of the recent descriptions of novel deep-branching archaeal lineages defined by SSU rDNA sequences recovered from psychrophilic environments (Hershberger *et al*. 1996; Jurgens *et al*. 1997; MacGregor *et al*. 1997; Schleper *et al*. 1997).

Compared with our knowledge of cultured Archaea, the surprising phylogeny and ecology of the marine Archaea Group I SSU rDNA sequences means that we await the cultivation of representative members with great expectation. This is also true for members of the env.pJP27/78 clade and those archaeal SSU rDNA sequences recovered in other ecological studies. The evidence suggests that the primary group Archaea underwent several early radiations resulting in genetically distant lineages. The present day members of these diverse archaeal lineages share an ecological range that is rapidly becoming equivalent to that of the other primary groups. Cultivation of representative species will allow us to uncover and explore these hidden biological resources.

J.O.M. and M.M. contributed equally to this work. This work was supported by the Forbairt Scientific Research Programme (SC/95/102). M.M. was supported by Forbairt and U.C.G. postgraduate fellowships. We are grateful to the captain and crew of the R.V. Prof. Marti and the staff of the U. C. G. Shellfish Research Unit, Carna. We thank T. Martin Embley, Mark Wilkinson, and the anonymous reviewers for constructive critique.

REFERENCES

Atlas, R. M., Colwell, R. R., Pramer, D., Tiedje, J. M., Vidaver, A. K. & Wodzinski, R, J. 1992 Microbial ecology. *ASM News* **58**, 45.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. 1992 *Short protocols in molecular biology*, 2nd edn. New York: Wiley.

Balch, W. E., Magrum, I. J., Fox, G. E., Wolfe, R. S. & Woese, C. R. 1977 An ancient divergence among the bacteria. *J. Molec. Evol.* **9**, 305–311.

Barns, S. M., Fundyga, R. E., Jeffries, M. W. & Pace, N. R. 1994 Remarkable Archaeal diversity detected in a Yellowstone National Park hot spring environment. *Proc. Natn. Acad. Sci. USA* **91**, 1609–1613.

Barns, S. M., Delwiche, C. F., Palmer, J. D. & Pace, N. R. 1996 Perspectives on Archaeal diversity, thermophily and

monophyly from environmental rRNA sequences. *Proc. Natn. Acad. Sci. USA* **93**, 9188–9193.
Bremer, K. 1988 The limits of amino-acid sequence data in angiosperm phylogenetic reconstruction. *Evolution* **42**, 795–803.
DeLong, E. F. 1992 Archaea in coastal marine environments. *Proc. Natn. Acad. Sci. USA* **89**, 5685–5689.
DeLong, E. F., Ying Wu, K., Prezelin, B. B. & Jovine, R. V. M. 1994 High abundance of Archaea in Antarctic marine picoplankton. *Nature* **371**, 695–697.
Eriksson, T. 1996 AutoDecay 2.9.2. Distributed by author. Available from `http://www.botan.su.se/systematik/folk/torsten.html`.
Felsenstein, J. 1981 Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Molec. Evol.* **17**, 368–376.
Felsenstein, J. 1985 Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791.
Fuhrman, J. A., Comeau, D. E., Hagstrom, Å. & Chan, A. M. 1988 Extraction from natural planktonic microorganisms of DNA suitable for molecular biology studies. *Appl. Environ. Microbiol.* **54**, 1426–1429.
Fuhrman, J. A., McCallum, K. & Davis, A. A. 1992 Novel major archaebacterial group from marine plankton. *Nature* **356**, 148–149.
Galtier, N. & Gouy, M. 1995 Inferring phylogenies from DNA sequences of unequal base compositions. *Proc. Natn. Acad. Sci. USA* **92**, 11317–11321.
Hershberger, K. L., Barnes, S. M., Reysenbach, A.-L., Dawson, S. C. & Pace, N. R. 1996 Wide diversity of Crenarchaeota. *Nature* **384**, 420.
Jurgens, G., Lindstrom, K. & Saano, A. 1997 Novel group within the Kingdom *Crenarchaeota* from boreal forest soil. *Appl. Environ. Microbiol.* **63**, 803–805.
Kishino, H. & Hasegawa, M. 1989 Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J. Molec. Evol.* **29**, 170–179.
Lane, D. J. 1991 16S/23S rRNA sequencing. In *Nucleic acid techniques in bacterial systematics* (ed. E. Stackebrandt & M. Goodfellow), pp. 115–175. New York: Wiley.
Lockhart, P. J., Steel, M. A., Hendy, M. D. & Penny, D. 1994 Recovering evolutionary trees under a more realistic model of sequence evolution. *Molec. Biol. Evol.* **11**, 605–612.
MacGregor, B. J., Moser, D. P., Wheeler Alm, E., Nealson, K. H. & Stahl, D. A. 1997 Crenarchaeota in Lake Michigan sediment. *Appl. Environ. Microbiol.* **63**, 1178–1181.
Maddison, W. P. & Maddison, D. R. 1992 MacClade, version 3. Sunderland, MA: Sinauer Associates Inc.
McInerney, J. O., Wilkinson, M., Patching, J. W., Embley, T. M. & Powell, R. 1995 Recovery and phylogenetic analysis of novel Archaeal rRNA sequences from a deep-sea deposit feeder. *Appl. Environ. Microbiol.* **61**, 1646–1648.
Olsen, G. J., Larsen, N. & Woese, C. R. 1991 The Ribosomal RNA Database Project. *Nucl. Acids Res.* **19**, 2017–2018.
Olsen, G. J., Matsuda, H., Hagstrom, R. & Overbeek, R. 1994 fastDNAml: a tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. *Comput. Appl. Biosci.* **10**, 41–48.
Preston, C. M., Ying Wu, K., Molinski, T. F. & DeLong, E. F. 1996 A psychrophilic crenarchaeon inhabits a marine sponge: *Cenarchaeum symbiosum* gen. nov., sp. nov. *Proc. Natn. Acad. Sci. USA* **93**, 6241–6246.
Raskin, L., Stromley, J. M., Rittmann, B. E. & Stahl, D. A. 1994 Group-specific 16S rRNA hybridization probes to describe natural communities of methanogens. *Appl. Environ. Microbiol.* **60**, 1232–1240.
Saitou, N. & Nei, M. 1987 The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molec. Biol. Evol.* **4**, 406–425.
Schleper, C., Holben, W. & Klenk, H.-P. 1997 Recovery of crenarchaeotal ribosomal DNA sequences from freshwater-lake sediments. *Appl. Environ. Microbiol.* **63**, 321–323.
Smith, S. 1994 GDE 2.2: the genetic data environment. Available from the Ribosomal Database Project server at `(http://rdp.life.uiuc.edu/)`.
Swofford, D. L. 1993 PAUP: phylogenetic analysis using parsimony. Distributed by author. Smithsonian Institute, Washington D.C.
Ueda, T., Suga, Y. & Matsuguchi, T. 1995 Molecular phylogenetic analysis of a soil microbial community in a soyabean field. *Eur. J. Soil Sci.* **46**, 415–421.
Watson, R. T., Heywood, V. H., Baste, I., Dias, B., Gamez, R., Janetos, T., Reid, E. & Ruark, G. 1995 *Global biodiversity assessment.* Cambridge/New York: University Press.
Woese, C. R. & Fox, G. E. 1977 Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natn. Acad. Sci. USA* **74**, 5088–5090.
Woese, C. R., Kandler, O. & Wheelis, M. L. 1990 Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natn. Acad. Sci. USA* **87**, 4576–4579.

Received 17 July 1997; accepted 8 August 1997