Diversity and genomics of Antarctic marine micro-organisms

Alison E. Murray* and Joseph J. Grzymski

Desert Research Institute, 2215 Raggio Parkway, Reno, NV 89512, USA

Marine bacterioplanktons are thought to play a vital role in Southern Ocean ecology and ecosystem function, as they do in other ocean systems. However, our understanding of phylogenetic diversity, genome-enabled capabilities and specific adaptations to this persistently cold environment is limited. Bacterioplankton community composition shifts significantly over the annual cycle as sea ice melts and phytoplankton bloom. Microbial diversity in sea ice is better known than that of the plankton, where culture collections do not appear to represent organisms detected with molecular surveys. Broad phylogenetic groupings of Antarctic bacterioplankton such as the marine group I Crenarchaeota, α-Proteobacteria (Roseobacter-related and SAR-11 clusters), γ-Proteobacteria (both cultivated and uncultivated groups) and Bacteriodetes-affiliated organisms in Southern Ocean waters are in common with other ocean systems. Antarctic SSU rRNA gene phylotypes are typically affiliated with other polar sequences. Some species such as Polaribacter irgensii and currently uncultivated γ-Proteobacteria (Ant4D3 and Ant10A4) may flourish in Antarctic waters, though further studies are needed to address diversity on a larger scale. Insights from initial genomics studies on both cultivated organisms and genomes accessed through shotgun cloning of environmental samples suggest that there are many unique features of these organisms that facilitate survival in high-latitude, persistently cold environments.

Keywords: Antarctic bacterioplankton; marine bacteria; marine microbial genomics; bacterioplankton diversity; Southern Ocean

1. INTRODUCTION

Polar marine environments are unique. They provide cold, saline water to fuel the deep-ocean conveyor belt (Baines & Condie 1998) and pump inorganic and organic carbon into the sequestered deep ocean (Sarmiento & Toggweiler 1984; Honjo 2004). Like the ocean below 200 m, the Southern Ocean is permanently cold with temperatures consistently below 4°C (Broecker & Peng 1982). In the majority of the Antarctic, seawater is at −1.8°C constantly. However, unlike the deep ocean, polar marine environments undergo extreme seasonal variations in sea-ice cover, light levels and day length. This greatly affects the biology and results in high rates of primary production and a strong pulse of biomass in summer (Holm-Hansen et al. 1977; Moline & Prezelin 1996) followed by a long winter of very low rates of photoautotrophy. Polar environments have been shown to be sensitive to global warming (Loeb et al. 1997; Smith et al. 2003; Moline et al. 2004), where minimal temperature differences can have significant effects on the extent and thickness of sea ice (Smetacek & Nicol 2005). Our knowledge of the concomitant dynamics and diversity of Antarctic marine bacterioplankton is still in its infancy. Only an adequate understanding of microbial diversity and genome-enabled capabilities in polar environments will enable us to complete the picture of polar ecosystem structure and function as well as establish the effects of climate change on the base of the polar food chain.

Observations of Antarctic bacterioplankton by light microscopy occurred at least as early as the Challenger expedition (1872–1876). These observations would have been limited to the larger bacterial species and predominantly phytoplankton. Cultivation-based studies of the domain bacteria appeared in journals almost 100 years later (Kriss et al. 1967) and have continued (Irgens et al. 1989; Delille 1996; Bowman et al. 1998; Bowman 2001) reporting cultivation of both cosmopolitan and endemic heterotrophic bacteria. However, it has been shown that cultivation-based studies overwhelmingly fail to describe either the microbial diversity of a given environment or the most active groups of organisms in that environment (Giovannoni 1991; Fuhrman et al. 1993; Amann et al. 1995). An exception may be the sea-ice bacterial community where cultivability well exceeds 1% of the viable population of bacteria (Brinkmeyer et al. 2003). And so, over the last decade, efforts to describe microbial diversity and activity in marine Antarctic environments have focused on molecular-based work.

This paper is an overview of molecular diversity and genomic research concerning Antarctic marine bacterioplankton. It focuses both on molecular applications making use of SSU rRNA gene diversity, fluorescence in situ hybridization (FISH) and real-time PCR, and on environmental genomics and genome sequencing of Antarctic micro-organisms. We have presented the

* Author for correspondence (alison.murray@dri.edu).

One contribution of 10 to a Theme Issue 'Antarctic ecology: from genes to ecosystems. Part 2: evolution, diversity and function'.

Published online 21 May 2007

doi:10.1098/rstb.2006.1944

Phil. Trans. R. Soc. B (2007) 362, 2259–2271

2259 This journal is © 2007 The Royal Society...
results of recent research and integrated them with other studies that have focused on the diversity and genomics of micro-organisms in Antarctic surface waters. This environment should be differentiated from the cold, deep-water environments that make up 70% of the ocean by volume. While organisms in these two environments are at constant cold temperatures, their ecology is controlled by vastly different physical, chemical and biological variables. However, in discussing the advances in Antarctic microbiology through the use of modern molecular techniques, it is impossible to ignore some of the major milestones in microbial genomics that have occurred in Antarctic deep water (Moreira et al. 2004), lakes (Saunders et al. 2003), soils and other cold habitats on Earth (Rabus et al. 2004).

2. MATERIAL AND METHODS
(a) **PCR-DGGE**
Seawater samples were collected on monthly intervals during both 1996–1997 and 2001–2002. DNA extraction and PCR-DGGE methods reported previously for the 1996–1997 samples (Murray et al. 1998) were used for the 2001–2002 samples except that samples (approx. 800 ng each) were run through 8% acrylamide, 1 mm thick, 30–60% denaturing gradient gels (62 V, 16 h) on a Bio-Rad D-Code system. Samples were stained for 15 min in SYBR-Green (Molecular Probes, Inc.) and then photographed (EDAS 290, Kodak). Here, we reanalysed the 1996–1997 data in comparison with the 2001–2002 dataset using Gel Compar II (Applied Maths, Inc.). Bands less than 1% of the total lane intensity were not considered. Pairwise comparisons using Sorenson’s index (Magurran 1988), which was used originally to compare the Antarctic bacterioplankton DGGE profiles (Murray et al. 1998), were calculated for all months in comparison to the August (late winter) sample.

(b) **Marine bacterial SSU rRNA gene diversity**
Bacterioplankton diversity was surveyed using three approaches (bacterial PCR-amplified SSU rRNA gene clone library, shotgun fosmid clone library and culture collection) and included a SSU rRNA gene clone library from sea ice for comparison. The planktonic sample for the SSU rRNA gene library was collected in spring (designated AntCL1 and AntCL2; 17 October 2001, 20 m off LTER station A); the large-contig shotgun library was prepared from a sample collected in late winter 1996 (designated Ant; for details on library see Béjà et al. 2002a); and samples for the bacterial culture collection were collected in October 2001 from seawater and sea ice. Sea-ice cores were collected (6 October and 21 October 2001) using a SIPRE corer in the Arthur harbour region (300 m off Palmer Station: clones AntCL3A1–D12 and 100 m off Palmer Station: clones AntCL3E1–H12).

A standard approach was used here for SSU rRNA gene library creation and sequencing. Briefly, universal primers targeting bacteria (27F and 1492R) were used to amplify SSU rRNA genes from bacterioplankton DNA extracted from samples collected on 0.22 μm filters following prefiltration (GF/A filters, Whatman, Inc.). The amplification products were ligated into plasmid vectors and transformed into Escherichia coli cells following manufacturer’s instructions (TOPO TA, Invitrogen, Inc.). Clones containing inserts were sequenced unidirectionally (T7 vector primer); then in most cases, clones with unique sequences were fully sequenced using the T3 vector primer and Bact533F. Sequences were run through CHECK_CHIMERA (Maidak et al. 1997) and BELLEROPHON (Huber et al. 2004) to identify potentially chimeric sequences.

Marine bacteria were cultivated using standard microbiological methods on several kinds of liquid and agar media at 4, 10 and 20°C. Media used were as follows (abbreviations were also used in strain IDs): SW media were prepared using coastal seawater and no other amendments; YE media included seawater amended with 5 mg l$^{-1}$ yeast extract; SC media were used to enrich Actinobacteria (Adelaye et al. 2004); DMSP media were used to enrich DMSP-using micro-organisms (Gonzalez et al. 1999); and NZ media included seawater amended with 5 mg l$^{-1}$ NZ-amine (Suna-Aldrich, Inc.). Isolates were initially selected based on phenotypic characteristics and then screened using PCR-DGGE to eliminate redundancy as described above. Full-length SSU rRNA gene sequences were determined for isolates with unique melting points.

SSU rRNA gene sequences were aligned with neighbouring sequences identified using BLAST (Altschul et al. 1990) and SEQUENCE MATCH (Cole et al. 2005). Sequences and alignment tools at the ribosomal database project were used for multiple sequence alignment (Cole et al. 2005). Phylogenetic relationships were determined with neighbour-joining method available in MEGA v. 3.1 (Kumar et al. 2004). Sequences submitted to GenBank have the following accession numbers: DQ06709–DQ906772 and DQ92580–DQ925858.

Details describing the shotgun library bacterial composition, gene content and amino acid modification analysis of Antarctic fosmid sequences were recently described by Gryzmski et al. (2006).

3. ECOLOGY AND DIVERSITY OF ANTARCTIC MARINE ARCHAEA
The first molecular ecological investigation in Antarctic waters targeting the bacterioplankton reported high abundance of archaeal RNA picoplanктon smaller than 1 μm in Antarctic surface waters (20–25% of total RNA) sampled in late winter/early spring (DeLong et al. 1994). The Antarctic SSU rRNA gene sequences corresponded to marine group I Crenarchaeota and marine group II Euryarchaeota that were just being reported at that time from temperate latitudes in the Pacific and Atlantic Oceans (DeLong 1992; Fuhrman et al. 1992; DeLong et al. 1994). This study helped to dispel the paradigm that marine group I Crenarchaeota were thermophiles by showing that these organisms exist in high abundance in coastal Antarctic surface waters. Subsequent studies investigating the temporal distribution of the marine group I Crenarchaeota showed that they appeared to vary substantially in per cent composition over the annual cycle in surface waters with higher abundance in winter as per cent of total RNA (Murray et al. 1998) or per cent of total cell counts (Church et al. 2003). Archaeal abundance also increased with depth, where cell concentrations were observed to vary between winter and summer (Church et al. 2003). Further studies demonstrated the circumpolar distribution of the planktonic archaea (Murray et al. 1999), as well as its variation over short time frames in coastal vertical profiles (Massana et al. 1998).

The diversity of both the Antarctic Crenarchaeota and Euryarchaeota appears to be quite limited in

*Phil. Trans. R. Soc. B* (2007)
comparison to marine bacterial phylogenetic groups (Massana et al. 2001). Crenarchaeal diversity has been reported to increase with depth and to be quite similar between assemblages sampled at both poles (Bano et al. 2004). The implications of this are not well understood, as it is possible that the SSU rRNA gene sequence variation under-represents genomic heterogeneity (Béja et al. 2002a). The SSU rRNA gene is a very slowly evolving molecule with an estimated 1% sequence variation accumulating over 50 Myr (Ochman & Wilson 1987); interestingly, this coincides roughly with the timing of the formation of the Antarctic circumpolar current which allowed the cooling of the Antarctic continent. It is possible that the marine group I Crenarchaeota, which appear to be global residents of cold waters and are the most abundant marine prokaryote in the world oceans (Karner et al. 2001), have undergone adaptation and diversification at the genome level to facilitate their survival and persistence in cold waters of the Antarctic and deep sea.

Until recently, the biogeochemical role of the marine group I Crenarchaeota has only been hypothesized given the difficulty in detecting archaeal metabolic activity in situ (Massana et al. 1998). However, recently, a representative marine group I-affiliated crenarchaeon capable of chemoheterotrophic growth using ammonia as an electron donor was isolated (Konneke et al. 2005). Subsequently, Francis et al. (2005) and Wuchter et al. (2006) reported on the diversity and distribution of archaeal ammonia mono-oxygenase genes in marine environments, supporting the hypothesis that ammonia-oxidizing Crenarchaeota are widespread. Another study has recently reported both $^{3}H$-leucine and $^{14}C$-bicarbonate uptake specifically by Crenarchaeota and Euryarchaeota in Antarctic circumpolar waters sampled between 200 and 3000 m (Herndl et al. 2005). The authors concluded that the Archaea may provide a significant CO$_2$ sink in the ocean ecosystem, particularly deep waters, and they speculated that organic carbon (the proxy in their study was an amino acid) might be used as an energy source. Thus, the extent of metabolic capabilities of the marine group I Crenarchaeota and relatives in marine sediments are currently unknown. Future studies promise to open this door further, particularly in Antarctic waters, as another recent study has shown that Crenarchaeota in marine sediments appear to grow heterotrophically (Biddle et al. 2006).

4. ECOLOGY AND DIVERSITY OF ANTARCTIC MARINE BACTERIA

The bacterial lineage Cytophaga–Flavobacteria–Bacteroides (CFB) has been a strong focus of both culture dependent (Bowman et al. 1997; Staley & Gosink 1999; Junge et al. 2002) and independent (Brinkmeyer et al. 2003; Abell & Bowman 2005) studies in part owing to their apparent prevalence in Antarctic marine waters and sea ice (Staley & Gosink 1999). Molecular-based studies using PCR–DGGE approaches (Webster et al. 2004; Abell & Bowman 2005), clone libraries (Brinkmeyer et al. 2003), FISH (Brinkmeyer et al. 2003; Abell & Bowman 2005) and real-time PCR (Abell & Bowman 2005) have all shown CFB to be an important constituent of sea-ice microbial communities (SIMCO) and Antarctic marine waters. The Flavobacteria, in particular, are a group suspected to thrive in high concentration of dissolved organic matter in sea ice (Brinkmeyer et al. 2003) and in algae blooms where they constitute as much as 70% of the bacteria (Glöckner et al. 1999; Abell & Bowman 2005). The abundance of particle-associated Flavobacteria correlates well with chl a concentration (Abell & Bowman 2005) and during periods of high productivity. In contrast, marine group I Crenarchaeota were inversely correlated to chl a concentration (Murray et al. 1998).

Molecular ecological studies on the diversity and ecology of marine bacterioplankton are not limited to Archaea and the CFB, though most of the reported diversity from the plankton arises from cultivation-based studies (Maugeri et al. 1996; Bowman et al. 1997; Michaud et al. 2004) and a recent SSU rRNA and rDNA sequencing survey (Gentile et al. 2006). Psychrophilic sea-ice bacteria ($\gamma$-proteobacterial and CFB-related strains) were found to be quite distinct from those in underlying waters, whereas the psychrotrophic strains were largely in the same phylogenetic groups (Bowman et al. 1997). Planktonic surveys performed off the Antarctic Italian base at Terra Nova Bay reported greater diversity of rRNA-encoded gene sequences than for DNA sequences (Gentile et al. 2006) with the greatest diversity in the $\gamma$-proteobacterial and CFB lineages. Comparisons of the results from two studies, one investigating diversity of bacterioplankton cultivars in Terra Nova Bay through comparisons of RFLP patterns (146 isolates were categorized into 52 distinct groups; Michaud et al. 2004) and another which compared RFLP patterns from a SSU rRNA gene clone library from McMurdo Sound seawater (160 clones fell into 103 groups; Webster et al. 2004), suggest that the diversity in the uncultivated fraction of the Antarctic bacterioplankton is relatively high and is comparable to other marine ecosystems.

Using PCR–DGGE, Murray et al. (1998) reported significant shifts of bacterioplankton communities as the seasons changed. These shifts were accompanied by significant decreases of marine group I Crenarchaeota in the water during periods of high biomass. Here, we have analysed a second set of bacterioplankton samples over the annual cycle and found that bacterial SSU rRNA gene variation in DGGE profiles compared over the annual cycle for two different years, 6 years apart, indicates that the variation over the annual cycle is consistent and the timing of shifts in assemblage composition is surprisingly reproducible (figure 1). Similarity of PCR–DGGE profiles calculated using Sorensen’s index suggests that the community is drastically different between winter and summer (figure 1; less than 45% similarity in both years).

Though bacterial SSU rRNA gene clone library sequencing of Antarctic bacterioplankton is relatively under-explored, a few studies investigating SIMCO diversity (Brown & Bowman 2001; Brinkmeyer et al. 2003) have found that many environmental sequences are closely related to those found in SIMCO culture collections—this is unusual, and in contrast to nearly
The phylogenetic relationships between the three most abundant classes of bacteria were determined to get a more detailed look at the planktonic diversity within these groups. The dominant α-proteobacterial sequences branched closely with *Pelagibacter ubique* (Rappé et al. 2002; Giovannoni et al. 2005), and other SAR-11-related bacteria from Arctic waters (figure 3). The α-proteobacterial sequences we detected were most closely related to sequences sampled in other polar studies (Bano & Hollibaugh 2002; Brinkmeyer et al. 2003; Arrieta et al. 2004). We detected several *Roseobacter* and *Sulfutobacter*-related sequences; however, we did not detect any sequences that were highly related to *Octadecabacter antarcticus*, a common genus in sea ice and seawater that contains organisms capable of producing gas vesicles. Several sequences (AntCL1B9, AntCL1H2 and AntCL1E1) branching off the *Octadecabacter* group (95% related) grouped tightly with the RCA cluster known only through environmental surveys recently described to be very abundant in the Southern Ocean (Selje et al. 2004).

Organisms affiliated with γ-Proteobacteria, the most cosmopolitan marine planktonic bacteria, were well represented in both environmental surveys and our culture collections though there was no overlap in species detection. The most abundant SSU rRNA gene sequences in the environmental PCR and shotgun clone libraries grouped with two distinct clusters that contain no known cultivated representatives (figure 4). The Ant10A4-related sequences branch off a cluster of sulphur-oxidizing symbionts and the closest cultivated member of the Ant4D3 group, which appears to be common at both poles, is most closely affiliated with an isolate in the Oligotrophic Marine Group (HTTC2180) at 91% SSU rRNA gene identity (Cho & Giovannoni 2004). We fully sequenced environmental fosmid clones affiliated with each to get more information concerning these little known groups (Ant4D3, Grzymski et al. 2006, and Ant10A4).

Our culture collection was dominated with a diverse group of γ-Proteobacteria mostly limited to a region of the tree with other well-studied organisms (e.g. *Pseudoalteromonas antarctica*, *Vibrio fischeri*, *Alteromonas macleodi*). Consistent with the findings of Brinkmeyer et al. (2003), several sea-ice environmental SSU rRNA sequences (AntCL3F9, AntCL3G12 and AntCL3E3) were most closely related to cultured organisms (e.g. *Glaciecola pallidula*, *Colwellia maris*).

The majority of Bacteriodetes sequences we detected were affiliated with the *Flavobacteriaceae* family, one of the most common groups in marine (Kirchman 2002; Kirchman et al. 2003) and polar (Abell & Bowman 2005) ecosystems. In particular, we detected a high number of *Polaribacter* irgensii-related sequences (figure 5). Unlike the dominant γ-Proteobacteria groups which were only represented by uncultivated members, the *Polaribacter* genus is represented by both cultivated and environmentally detected sequences. *Polaribacter* irgensii-related sequences, though not confined to polar environments, are closest to organisms and environmental sequences from both Arctic and Antarctic environments, though they are thought to be more highly represented in the Antarctic (Brinkmeyer et al. 2003) and were one of the

---

*Phil. Trans. R. Soc. B* (2007)
most common Antarctic zone groups detected in a latitudinal survey (Abell & Bowman 2005). We also detected a number of sequences grouping in clusters known only from environmental surveys, many in the Flavobacteriace (Abell & Bowman 2005). The most abundant of these grouped in the Agg58/Cryomorphaceae cluster and were most similar to uncultivated sequences from the Arctic (Arctic97A-17; Bano & Hollibaugh 2002). Overall, the Bacteriodetes diversity detected, particularly in the free-living planktonic library, was high, consistent with the suspected importance of this group in polar ocean systems where they may be particularly associated with algal blooms (Simon et al. 1999) and particles (Abell & Bowman 2005).

5. ANTARCTIC ENVIRONMENTAL GENOMICS

The environmental genomics approach, which has recently become popular in part due to the lower cost of sequencing (Tyson et al. 2004; Venter et al. 2004), has well-established roots in Antarctic bacterioplankton studies albeit on a smaller scale (Béjà et al. 2002; de la Torre et al. 2003; Lopez-Garcia et al. 2004; Moreira et al. 2004, 2006; Grzymski et al. 2006). In fact, Antarctic environmental genomics has perhaps outpaced complete microbial genome sequencing until very recently. Large DNA inserts (30–45 kb) provide more robust information than ribosomal RNA gene phylogenetic diversity, including gene content and biological potential, which is particularly valuable for organisms that largely remain uncultivable (Béjà et al. 2002a).

These larger genome fragment datasets provided information about genomic divergence that was otherwise lost when comparing SSU rRNA gene sequences from highly similar phylotypes. In a study comparing marine group I crenarchaeotal genome fragments from Antarctica and Pacific temperate waters, strain variability was shown to be considerably different in four Antarctic Crenarchaeota despite almost identical SSU rRNA gene sequences (greater than 99.8% identical; Béjà et al. 2002a). Furthermore, these data established chromosomal organization in marine group I Crenarchaeota despite there being no cultivated representatives at that time. While protein prediction only suggests functional diversity, this greatly expands our understanding of microbial ecology and the metabolic potential of uncultivable organisms (Béjà et al. 2002a; de la Torre et al. 2003; Moreira et al. 2004; Grzymski et al. 2006). So far this has best been shown in the pervasiveness of proteorhodopsin genes in bacteria including the SAR-86, γ-Proteobacteria and ubiquitous SAR-11 cluster of α-Proteobacteria (Béjà et al. 2001, 2002b; de la Torre et al. 2003; Venter et al. 2004; Giovannoni et al. 2005), all of which have been detected in environmental genomic surveys. These studies extended the use of environmental genomics to describe functional gene diversity.

Since studies of functional gene diversity (i.e. nirK, RuBiSCO, AmoA) often preclude phylogenetic classification of the sequences, either the inclusion of highly conserved genes in the environmental fragment or the preponderance of evidence from BLAST results can be used to identify the affiliation of the genome fragment. de la Torre et al. (2003) used 13 ribosomal proteins and BLAST results of the remaining predicted proteins to conclude that a proteorhodopsin-containing BAC from the central North Pacific was...
most probably affiliated with an \(\alpha\)-proteobacterium. Less definitively, the PR-containing Antarctic fragment and another highly syntenous fragment (96% nucleotide identity over 20 kbp) could only be identified as Proteobacteria affiliates (36% of predicted proteins most similar to \(\gamma\)-Proteobacteria).

Similarly, the Antarctic fosmid clone Ant24C4 was identified as an \(\alpha\)-Proteobacteria from BLAST results of its 42 ORFs and another fosmid clone Ant29B7 fragment was affiliated with the Bacteroidetes phylum from phylogenetic analysis of \textit{groEL} (Grzymski \textit{et al.} 2006).
Figure 4. Neighbour-joining tree showing phylogenetic relationships between Antarctic marine plankton and sea-ice γ-Proteobacteria and many of their nearest neighbours isolated or directly sequenced from marine environments. Symbols are as described in figure 3 but shotgun clone libraries are indicated by open circles and in addition, open squares represent sequences for cultivated organisms and nodes representing both SSU rRNA gene libraries.

Comparative analyses within highly similar populations and between distant ecosystems are necessary to understand the ecology and evolution of marine bacterioplankton including the effects of gene loss, rearrangement and horizontal gene transfer. The comparative analysis of Antarctic marine crenarchaeotal genome fragments has revealed multiple horizontal gene transfers (Lopez-Garcia et al. 2004). These
horizontal gene transfers are hypothesized to play a role in the low-temperature adaptation of marine group I Crenarchaeota. The complex evolutionary history of some of these genes was demonstrated by the phylogenetic analysis of ORFs surrounding the conserved 16S-23S-ITS-GSAT region of Crenarchaeota. The same large insert library also resulted in the recovery of the first group II Euryarchaeota genome fragment to contain the 16S rRNA gene (Moreira et al. 2004). Incidentally, environmental genomic surveys in the North Pacific subtropical gyre recently found group II Euryarchaeota-encoded proteorhodopsin genes in surface water samples (Frigaard et al. 2006), consistent with the prevalence of these organisms in the photic

Figure 5. Neighbour-joining tree showing phylogenetic relationships between Antarctic marine plankton and sea-ice Bacteriodetes-related sequences and many of their nearest neighbours isolated or directly sequenced from marine environments. Symbols are the same as in figures 3 and 4.
zone (DeLong 1992; Massana et al. 1997; DeLong et al. 2006). The genome fragment analysis of ORFs perhaps revealed a novel succinate dehydrogenase (Moreira et al. 2004). Again, the work emphasized the use of genome fragments to infer metabolic potential in the uncultivable organisms. Metabolic potential must be emphasized until more data are available (e.g. expression studies, metabolic assays).

Recently, six environmental genome fragments from nearshore Antarctic coastal waters were described (Grzymski et al. 2006). These bacterial fragments from a late winter library (Béja et al. 2002a) were chosen for complete sequencing based on their hypothesized ecological relevance (one γ- Proteobacteria, one putative α-Proteobacteria, two Bacteroidetes) and uniqueness (one Gemmatimonadetes, one Actinobacteria). The genome fragments that contained SSU rRNA genes were only 85–91% similar to 16S rRNA genes from cultivated relatives. Apart from describing the content of the six fragments, the authors also analysed the predicted proteins against the SwissProt database and found amino acid usage patterns indicative of cold adaptation. While this analysis has been performed on various psychrophile enzymes such as α-amylase and β-glucosidase (Feller & Gerday 1997; Huston et al. 2004) and whole genomes (Saunders et al. 2003; Methé et al. 2005), this was the first analysis conducted across diverse bacterial taxa. The most significant amino acid changes in the predicted proteins were reductions in salt-bridge-forming residues (arginine), reduced proline content and reduced hydrophobic clusters (Figure 6; Grzymski et al. 2006). In addition, predictions of protein disordered regions were significantly longer in Antarctic sequences than in mesophile homologues. The lengthened disordered regions (regions without a defined three-dimensional structure) are hypothesized to aid in processes that are retarded at cold temperature such as various association and disassociation reactions (Dunker et al. 2002; Grzymski et al. 2006).

### 6. GENOME SEQUENCING OF POLAR MICRO-ORGANISMS

As of spring 2006, there are nine bacterial and one archaean polar microbial genomes in GenBank. Several are accompanied by publications: two Euryarchaeota (Methanogenium frigidum and Methanococcoides burtonii; Saunders et al. 2003), two γ-Proteobacteria (Colwellia psychrerythraea 34H Methé et al. 2005; Pseudoalteromonas haloplanktis, Medigue et al. 2005) and a δ-proteobacterium (Desulfitalea psychrophila; Rabus et al. 2004). At least nine other polar marine strains are in various stages of completion (Table 1) in addition to at least ten other psychrophilic strains including fish pathogens and deep-sea strains. Seven out of the 16 polar bacteria are from Antarctic surface waters or sea ice: two CFBs, Psychroflexus torquis (isolated from a sea-ice algal assemblage), P. irgensii (isolated from nearshore waters off the Antarctic Peninsula), a Marine Actinobacterium PHSC20C1 (also from Antarctic Peninsula waters), O. antarcticus 307, an α-proteobacterium strain isolated from McMurdo Sound, and γ-Proteobacteria strains, P. haloplanktis, Oleïspira antarctica and Shewanella frigidimarina, which were isolated from either seawater or sea ice.

The CFB genomes represent two of the three sequences available from polar marine habitats and promise to offer much in the way of understanding features that facilitate their dominance in these ecosystems. For example, initial investigations have revealed ice-binding and cold-shock proteins suggesting specialization to cold environments (Table 2). The bipolar distribution of P. irgensii (Staley & Gosink 1999), its prevalence in sea ice and seawater, and cellular features such as a proteorhodopsin homologue and gas vesicle formation suggest that this organism will be a model for polar marine microbiology and cold adaptation studies in the future.

A study of the draft genome sequences of two Acid Lake methanogens, M. frigidum and M. burtonii, was the first comparative analysis of archaean genomes that spanned psychrophile to hyperthermophile lifestyles (Saunders et al. 2003). The analysis of the two Antarctic methanogens detailed genotype characteristics of cold adaptation such as increased glutamine and threonine usage and decreased leucine usage in the predicted proteins of these genomes. Five predicted
proteins were unique to the two organisms and have putative roles in adaptation to the cold. Three of these cold-shock domain folds. The comparative analysis of nine methanogen genomes spanning an optimal growth temperature (OGT) of 83°C also showed that G+C content is a major factor influencing tRNA stability in organisms with an OGT greater than 60°C. G+C content was not significantly different below that temperature (Saunders et al. 2003).

Similar comparative analyses were done on the Arctic psychrophile bacterium C. psychrerythraea 34H and 24 other organisms with wide-ranging OGTs (Methé et al. 2005). These data suggested a decrease in charged amino acids in the psychrophile when compared with mesophiles and thermophiles. It also appeared that serine usage increased as OGT decreased. In the P. haloplanktis genome, which appears to be well equipped for growth in the cold, a significant bias towards asparagine residues was found (Medigue et al. 2003).

7. CONCLUSIONS
In conclusion, the state of knowledge concerning polar marine micro-organisms from ecological and genomic perspectives is in the early phase of exponential growth. Initial studies on diversity of bacterioplankton suggest

Table 1. Polar bacterial and archaeal genome projects from both marine and terrestrial environments. (The projects without accession numbers have not been submitted to GenBank and for these cases, the URLs for genome projects are listed if available.)

<table>
<thead>
<tr>
<th>domain</th>
<th>phylogenetic group</th>
<th>micro-organism</th>
<th>origin of strain</th>
<th>status of genome/accession number or URL</th>
<th>reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Archaea</td>
<td>Euryarchaeota</td>
<td>Methanococcales burtonii</td>
<td>Ace Lake, Antarctica</td>
<td>draft/CP000300</td>
<td>Saunders et al. (2003)</td>
</tr>
<tr>
<td>Bacteria</td>
<td>γ-Protobacteria</td>
<td>Coeloeila psychrerythraea</td>
<td>Arctic Marine Sediments</td>
<td>complete/CP000083</td>
<td>Methé et al. (2005)</td>
</tr>
<tr>
<td>Bacteria</td>
<td>γ-Protobacteria</td>
<td>Stentamella frigidimarina</td>
<td>Sea Ice, Antarctica</td>
<td>complete/CP000082</td>
<td></td>
</tr>
<tr>
<td>Bacteria</td>
<td>γ-Protobacteria</td>
<td>Psychrobacter arcticus</td>
<td>Siberian permafrost</td>
<td>complete/CP000323, CP000324</td>
<td></td>
</tr>
<tr>
<td>Bacteria</td>
<td>γ-Protobacteria</td>
<td>Psychrobacter cryohalolentis</td>
<td>Siberian permafrost (40k yr old)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacteria</td>
<td>γ-Protobacteria</td>
<td>Oleispira antarctica</td>
<td>Rod Bay, Ross Sea, crude oil-enriched surface water</td>
<td>in progress</td>
<td></td>
</tr>
<tr>
<td>Bacteria</td>
<td>γ-Protobacteria</td>
<td>Pseudoalteromonas haloplanktis</td>
<td>Coastal Antarctic seawater, Terra Adelie</td>
<td>complete/CR954246, CR954247, CR522870, CR522871, CR522872</td>
<td>Medigue et al. (2005)</td>
</tr>
<tr>
<td>Bacteria</td>
<td>δ-Protobacteria</td>
<td>Desulfitotalea psychrophila</td>
<td>Arctic marine sediments, Svalbard</td>
<td>draft/AADW00000000, draft/AAPR00000000</td>
<td>Rabus et al. (2004)</td>
</tr>
<tr>
<td>Bacteria</td>
<td>Firmicutes Bacteroides</td>
<td>Exiguobacterium</td>
<td>Siberian permafrost sea-ice algal assemblage, Prydz Bay, Antarctica</td>
<td>draft/CP000084, CP000085</td>
<td></td>
</tr>
<tr>
<td>Bacteria</td>
<td>Bacteroidetes</td>
<td>Psychrosphexxus torquis</td>
<td>Arctic surface seawater, 350 km north of Deadhorse, Alaska</td>
<td>in progress</td>
<td></td>
</tr>
<tr>
<td>Bacteria</td>
<td>Bacteroidetes</td>
<td>Polaribacter filamentos 215</td>
<td>nearshore marine waters off Antarctic Peninsula</td>
<td>draft/AAOG00000000</td>
<td></td>
</tr>
<tr>
<td>Bacteria</td>
<td>Bacteroidetes</td>
<td>Polaribacter ignessis 23-P</td>
<td>nearshore marine waters of Antarctic Peninsula</td>
<td>draft/AAOB00000000</td>
<td></td>
</tr>
<tr>
<td>Bacteria</td>
<td>Actinobacteria</td>
<td>Marine Actinobacterium</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Examples of proteins involved in cold resistance in several polar microbial genomes (based on function in the environment, adaptation to cold and natural product synthesis). (All these proteins are present in the P. ignessis 23-P genome.)

<table>
<thead>
<tr>
<th>cold-resistant proteins</th>
<th>cold-shock (acclimation) protein (EAR12042, EAR12901 and EAR11908)</th>
<th>soluble acyl-ACP desaturase (EAR13799)</th>
<th>ATP-dependent RNA helicase, DEAD/DEAH box family protein (EAR12918 and EAR11919)</th>
<th>β-galactosidase, putative (EAR13130)</th>
<th>putative ribonuclease H1 (EAR13453)</th>
</tr>
</thead>
</table>

that the diversity appears to rival that found in other ocean systems, though many polar phylotypes harbour a distinct biogeographic signal (Pommier et al. 2005). This information is essential if we are to be able to understand the specific adaptations of organisms to life in the cold, describe the evolutionary history of psychrophiles, and, importantly, predict the impacts of climate change, such as reduction in sea ice and its impact on the carbon cycle in polar oceans. Understanding the regulation and roles of ice-binding proteins (in P. ignessis), dioxygenases (in P. haloplanktis,
Marine bacterioplankton assemblages.

This work was the result of efforts by B. J. Carter, who is graciously thanked for his assistance in this research. A. E. Kelley played a key role in field research support. Both Carter and Kelley developed and maintained the culture collection. We also thank the Ribosomal Database Project for computational support. Sequencing collaborations at Joint Genome Institute were instrumental in this research. RPSC is acknowledged for their assistance in sample collection in 1997 and 2002. NSF-OPP 0085435 grant to A.E.M. supported this work in part.

REFERENCES


Tyson, G. W. et al. 2004 Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* 428, 37–43. (doi:10.1038/nature02340)

